

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): MIZUTANI, et al.
Serial No.: Not yet assigned
Filed: July 7, 2003
Title: METHOD AND SYSTEM FOR CONTENT-ORIENTED
ROUTING IN A STORAGE-EMBEDDED NETWORK
Group: Not yet assigned

LETTER CLAIMING RIGHT OF PRIORITY

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

July 7, 2003

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s) the right of priority based on Japanese Patent Application No.(s) 2002-302443, filed October 17, 2002.

A certified copy of said Japanese Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP



Carl I. Brundidge
Registration No. 29,621

CIB/alb
Attachment
(703) 312-6600

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出願年月日

Date of Application:

2002年10月17日

出願番号

Application Number:

特願2002-302443

[ST.10/C]:

[JP2002-302443]

出願人

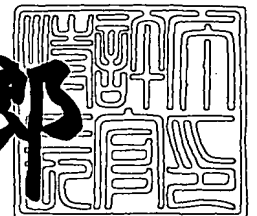
Applicant(s):

株式会社日立製作所

2003年 4月25日

特許庁長官
Commissioner,
Japan Patent Office

太田信一郎



出証番号 出証特2003-3029817

【書類名】 特許願

【整理番号】 H02015931A

【あて先】 特許庁長官 殿

【国際特許分類】 H04L 12/00

【発明者】

【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所中央研究所内

【氏名】 水谷 昌彦

【発明者】

【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所中央研究所内

【氏名】 池田 博樹

【発明者】

【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所中央研究所内

【氏名】 宮城 盛仁

【発明者】

【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所中央研究所内

【氏名】 木下 成顕

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社 日立製作所

【代理人】

【識別番号】 100075096

【弁理士】

【氏名又は名称】 作田 康夫

【電話番号】 03-3212-1111

【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 情報ルーティング方式および情報中継装置

【特許請求の範囲】

【請求項 1】

複数の入力ポートを持つ受信部と、複数の出力ポートを持つ送信部と、データ処理部と、前記受信部と送信部と前記データ処理部を接続するスイッチ部と、記憶機能を有する複数のストレージ部と、前記受信部、前記送信部、前記スイッチ部、前記ストレージ部を制御するルーティン処理部とを有するアプリケーション層でルーティング機能を持つ情報中継装置であって、

ネットワーク上で前記複数ストレージ部を識別するためのストレージアドレスと、前記ストレージアドレスと、データ属性と、を含むデータパケットを用いる情報中継手段を有し、

前記受信部は、データパケットを受信する手段と、前記データパケットからストレージアドレスと前記データ属性を取り出す手段と、前記データ属性をルーティング処理部に転送する手段と、前記データパケットを前記スイッチ部に送り込む手段とを有し、

前記ルーティング処理部は、ストレージアドレスを含む経路情報に基づいて受信データパケットの送出先を、前記送信部または前記データ処理部から選択し、前記スイッチ部に対して切替を命令する手段とを有し、前記ストレージ部は、前記受信データを記憶する手段とを有し、

前記スイッチ部は、前記制御信号に基づき経路の切替を行う手段を有し、

前記データ処理部は、データ属性に基づいて、データを格納もしくは送信する手段を有し、

前記送信部は、前記制御部からの制御信号に従いパケットのヘッダ処理を行う手段と、データパケットを隣接中継装置へ転送する手段と、を備えることを特徴とする情報中継装置。

【請求項 2】

前記情報中継装置は、ネットワーク上でストレージ部を識別する為の識別情報（ネットワークストレージアドレス：NSA）を用いて表される経路表（ストレージ

ルーティングテーブル：SRT) を制御部に保持する手段と、SRTを用いて当該宛先NSAに対応する送出先NSAを決定する手段と、を有することを特徴とする請求項1記載の情報中継装置。

【請求項3】

前記NSAは、中継装置のネットワーク上における位置情報と、物理的あるいは論理的に区別される記憶装置の個体識別情報と、保存領域内においてディレクトリまたはブロックアドレス指定によるデータ保存位置を識別する情報と、のうち一つまたは複数の組み合わせによって表されることを特徴とする請求項2記載の情報中継装置。

【請求項4】

前記情報中継装置は、ストレージ部にメモリあるいはHDDの一つ又は複数の媒体から構成される記憶領域を有することを特徴とする請求項1から3のいずれかに記載の情報中継装置。

【請求項5】

前期情報中継装置は、スイッチ部において受信データをストレージ部に保存する為、入力データパケットをデータ処理部に送る手段と、前期ストレージ部は前期データ処理部からデータを受け取り、少なくとも次段の中継装置への転送が完了するまで受信データを装置内に保存しておく手段を有することを特徴とする請求項1から4のいずれかに記載の情報中継装置。

【請求項6】

前記情報中継装置は、前記データ保存処理において、受信したデータをパケットの状態で記憶する手段と、複数のパケットからなるデータを装置内で再構成して記憶しておく手段とのいずれかまたは両方を有することを特徴とする請求項5記載の情報中継装置。

【請求項7】

前記情報中継装置は、受信部にてデータ送出要求を検出した場合に自装置内に記憶されているデータを取り出して再送する手段を有することを特徴とする請求項1から6のいずれかに記載の情報中継装置。

【請求項8】

前記情報中継装置は、SRTに登録する宛先NSAと次の中継用NSAとの対応を決定する際に、受信したデータフローのデータサイズと次の中継用ストレージ部における利用可能な記憶領域の容量とに基づいて経路を判断し、SRTを構成する手段を有することを特徴とする請求項1から7のいずれかに記載の情報中継装置。

【請求項9】

前記情報中継装置は、情報ルーティングの為にネットワークを構成する装置の間で相互に利用可能な記憶容量を通知する為の手段を有することを特徴とする請求項1から8のいずれかに記載の情報中継装置。

【請求項10】

前記情報中継装置は、受信部において、入力ポートに供給されたデータがストレージアドレスに基づいてルーティングされるべきものか否かを判定する手段を有することを特徴とする請求項1から9のいずれかに記載の情報中継装置。

【請求項11】

前記情報中継装置は、転送経路の決定に際して下位層の転送プロトコルの働きによって得られる経路情報を利用する手段を有することを特徴とする請求項1から10のいずれかに記載の情報中継装置。

【請求項12】

前期情報中継装置は、データの送信に際し、必要ならば、データをアプリケーション層において複数のパケットに分割して転送する手段を有することを特徴とする請求項1から11のいずれかに記載の情報中継装置。

【請求項13】

前期データパケットは、アプリケーション層のデータ属性を含むヘッダ部と、データの内容を含むデータ部とから構成されることを特徴とする請求項1から12のいずれかに記載の情報中継装置。

【請求項14】

前期データパケットは、ヘッダ部にデータの宛先NSAとデータの送信元NSAとを含むことを特徴とする請求項1から13のいずれかに記載の情報中継装置。

【請求項15】

前期データパケットは、ヘッダ部に含まれるデータ属性として、同一のデータか

ら生成されたパケットであることを示すデータ識別子と、データが分割されて送信される場合には再組み立て情報としてデータパケットの順序を示すパケット識別子とを含むことを特徴とする請求項 1 から 1 4 のいずれかに記載の情報中継装置。

【請求項 1 6】

前期データパケットは、データ属性により優先的に経路を決定する為、ヘッダ部にデータパケットの優先度情報を含むことを特徴とする請求項 1 から 1 5 のいずれかに記載の情報中継装置。

【請求項 1 7】

複数の入力ポートを持つ受信部と、複数の出力ポートを持つ送信部と、データ処理部と、前記受信部と送信部と前記データ処理部を接続するスイッチ部と、記憶機能を有する複数のストレージ部と、前記受信部、前記送信部、前記スイッチ部、前記ストレージ部を制御するルーティング処理部と、から構成されるアプリケーション層のデータ属性をもとにルーティングを行う情報ルーティング方式であって、

ネットワーク上で前記複数ストレージ部を識別するためのストレージアドレスと、前記ストレージアドレスと、データ属性と、を含むデータパケットを用いる情報ルーティング方式を特徴とし、

前記受信部は、データパケットを受信するステップと、前記データパケットからストレージアドレスと前記データ属性を取り出すステップと、前記データ属性をルーティング処理部に転送するステップと、前記データパケットを前記スイッチ部に送り込むステップとを有し、

前記ルーティング処理部は、ストレージアドレスを含む経路情報に基づいて受信データパケットの送出先を、前記送信部または前記データ処理部から選択し、前記スイッチ部に対して切替を命令するステップと、

前記ストレージ部は、前記受信データを記憶するステップとを有し、

前記スイッチ部は、前記制御信号に基づき経路の切替を行うステップを有し、

前記データ処理部は、データ属性に基づいて、データを格納もしくは送信するステップを有し、

前記送信部は、前記制御部からの制御信号に従いパケットのヘッダ処理を行うステップと、データパケットを隣接中継装置へ転送するステップと、
を備えることを特徴とする情報ルーティング方式。

【請求項 1 8】

前記制御ステップは、ネットワーク上でストレージ部を識別する為の識別情報（NSA）を用いて表される経路表（SRT）を制御部に保持する手段と、SRTを用いて当該宛先NSAに対応する送出先NSAを決定する手段と、を有することを特徴とする請求項 1 7 記載の情報ルーティング方式。

【請求項 1 9】

前記NSAは、中継装置のネットワーク上における位置情報と、物理的あるいは論理的に区別される記憶装置の個体識別情報と、保存領域内においてディレクトリまたはブロックアドレス指定によるデータ保存位置を識別する情報と、のうち一つまたは複数の組み合わせによって表されることを特徴とする請求項 1 8 記載の情報ルーティング方式。

【請求項 2 0】

前期制御ステップは、受信部において入力パケットのヘッダ処理を行うと共に受信データを制御部内メモリまたは記憶装置内に保存する為、入力データのコピーを制御部に送る手段と、制御部はヘッダ情報をもとにデータパケットを識別し、少なくとも次段のNSAへの転送が完了するまで受信したデータパケットの情報を装置内に保存しておく手段を有することを特徴とする請求項 1 7 から 1 9 のいずれかに記載の情報ルーティング方式。

【請求項 2 1】

前記制御ステップは、前記データ保存処理において、受信したデータをパケットの状態で記憶する手段と、複数のパケットからなるデータを装置内で再構成して記憶しておく手段とのいずれかまたは両方を有することを特徴とする請求項 1 7 から 2 0 のいずれかに記載の情報ルーティング方式。

【請求項 2 2】

前記制御ステップは、受信部にてデータ送出要求を検出した場合に自装置内に記憶されているデータを取り出して再送する手段を有することを特徴とする請求項

17から21のいずれかに記載の情報ルーティング方式。

【請求項23】

前記情報中継装置は、SRTに登録する宛先NSAと次の中継用NSAとの対応を決定する際に、受信したデータフローのデータサイズと次の中継用ストレージ部における利用可能な記憶領域の容量とに基づいて経路を判断し、SRTを構成する手段を有することを特徴とする請求項17から22のいずれかに記載の情報ルーティング方式。

【請求項24】

前記制御ステップは、情報ルーティングの為のネットワークを構成する装置の間で相互に利用可能な記憶容量を通知する為の手段を有することを特徴とする請求項17から23のいずれかに記載の情報ルーティング方式。

【請求項25】

前記制御ステップは、受信部において、入力ポートに供給されたデータがストレージアドレスに基づいてルーティングされるべきものか否かを判定する手段を有することを特徴とする請求項17から24のいずれかに記載の情報ルーティング方式。

【請求項26】

前記情報中継装置は、情報ルーティングの為のネットワークを構成する装置の間で相互に利用可能な記憶容量を通知する為のステップを有することを特徴とする請求項17から25のいずれかに記載の情報ルーティング方式。

【請求項27】

前記情報中継装置は、受信部において、入力ポートに供給されたデータがストレージアドレスに基づいてルーティングされるべきものか否かを判定するステップを有することを特徴とする請求項17から26のいずれかに記載の情報ルーティング方式。

【請求項28】

前記情報中継装置は、転送経路の決定に際して下位層の転送プロトコルの働きによって得られる経路情報を利用するステップを有することを特徴とする請求項17から27のいずれかに記載の情報ルーティング方式。

【請求項 2 9】

前期情報中継装置は、データの送信に際し、必要ならば、データをアプリケーション層において複数のパケットに分割して転送するステップを含むことを特徴とする請求項 1 7 から 2 8 のいずれかに記載の情報ルーティング方式。

【請求項 3 0】

前期データパケットは、アプリケーション層のデータ属性を含むヘッダ部と、データの内容を含むデータ部とから構成されることを特徴とする請求項 1 7 から 2 9 のいずれかに記載の情報ルーティング方式。

【請求項 3 1】

前期データパケットは、ヘッダ部にデータの宛先NSAとデータの送信元NSAとを含むことを特徴とする請求項 1 7 から 3 0 のいずれかに記載の情報ルーティング方式。

【請求項 3 2】

前期データパケットは、ヘッダ部に含まれるデータ属性として、同一のデータから生成されたパケットであることを示すデータ識別子と、データが分割されて送信される場合には再組み立て情報としてデータパケットの順序を示すパケット識別子とを含むことを特徴とする請求項 1 7 から 3 1 のいずれかに記載の情報ルーティング方式。

【請求項 3 3】

前期データパケットは、データ属性により優先的に経路を決定する為、ヘッダ部にデータパケットの優先度情報を含むことを特徴とする請求項 1 7 から 3 2 のいずれかに記載の情報ルーティング方式。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明はネットワーク経路制御方式及びネットワーク中継装置に係る。特に映像、音声等のマルチメディアデータおよび企業内技術データといった大容量データを効率的に転送する為、データパケットの持つアプリケーション層のデータ識別情報を用いてデータパケットのルーティングを行うデータ転送方式およびその中

継装置、またストレージ装置を構成要素とするネットワークにおいてストレージアドレスを用いるデータ転送方式とデータ蓄積型の転送方式、およびその中継装置に関する。

【 0 0 0 2 】

【従来の技術】

従来のネットワークシステムでは、ISO参照モデルのネットワーク層プロトコルによるルーティング技術が利用される。主なプロトコルとして、経路上に存在するルータの台数に基づき最適経路を決定するRouting Information Protocol(RIP)、リンク状態から優先経路を決定するOpen Shortest Path First(OSPF)、経路の属性に基づき自律システム間の優先通信経路を決定するBorder Gateway Protocol(BGP)が挙げられる。これらのプロトコルでは個々のルータが自装置内に保持されている経路情報を相互に通知し合うことによりネットワークトポロジが伝播され、各プロトコルに定められた計算方式により、ネットワーク層で定義されるアドレス情報および回線の種類等の情報に基づきデータの送信先と送出先インタフェースの対応表（ルーティングテーブル）が生成される。

ネットワークを形成するデータ中継装置（ルータ）は他のルータから学習した経路情報と自装置の持つ接続情報をもとに宛先アドレスと出力ポートの対応表をルーティングテーブルとして作成し、メモリ内に保持しておく。ルータ間ではルーティングプロトコルにより定期的に経路情報が交換され、接続の障害やネットワークトポロジの変化に対応して随時ルーティングテーブルが更新される。

ルーティングテーブル上には一つの宛先アドレスに対応する出力ポートとして、最も適当と判断されたものが一つ登録され、データパケットはこの経路表をもとに各ルータで送信先を決定されて、目的のネットワークアドレスをもつ中継装置または端末装置へと送られる。既存のIPネットワークでは、あるルータからある宛先に向かう経路はネットワークアドレスに基づいてただ一通りに決定されるという特徴があり、トラフィックがある区間に集中する場合においては負荷の高い区間と通常殆ど利用されない区間とに分かれ、ネットワーク帯域を有効に利用できない。

IPネットワークに対して機能を補うことで高速、セキュアな転送を行う技術が開

発されている。現在最も注目されているものは Multi Protocol Label Switching (MPLS)であり、IETF等でも標準化が進んでいる。MPLSではルータ間の転送においてラベルを用いたフロー識別を行い、トラフィック制御を実現する技術である。コネクションレスネットワークにおいてコネクション型の接続を確立するための手段であり、セキュリティ上の効果も高い。MPLSはネットワーク層のルーティングテーブルを参照して経路設定を行うが、トラフィックエンジニアリング技術により、帯域利用状況に応じたパスの確保を独自に行うことが可能である。RSVP-TEやCR-LDP等のプロトコルがIETFで標準化されており、ユーザ側のMPLS網エッジルータからデータ送信側のエッジルータに対し経路確立要求が送られる際に予約経路上のネットワーク利用状況に基づいて、より効率的な転送が可能な経路を設定する技術の検討は進められている。但し、経路を設定する時点でのネットワーク状態に応じた経路およびラベルの設定が行われる為、データ通信中の動的な経路変更には対応していない。

ネットワークの利用状況を集中管理することにより効率的な経路を設定するポリシールーティングと呼ばれる方式では、ネットワーク帯域の輻輳、障害といった情報に対して個々のルータがあらかじめ設定された動作判断（ポリシー）に基づいて経路を動的に変更することが可能である。但し、経路選択は有限個の経路に限られる。

また、大容量データ配信の為のネットワーク要素技術は以下のとおりである。データの大容量化と企業情報システムの拡大に伴い、コンテンツ配信網（CDN）と呼ばれるシステムの需要が拡大している。リクエストに対するレスポンスの向上により、データの可用性を高めることを目的とする。具体的には、データの分散配置によりサーバ別のアクセス数を制限しパフォーマンスを向上させること、データ転送時の遅延を低減することが主な課題とされている。

データをエンドユーザへ効率的に配信する手段としてキャッシュをアクセス網側に配置しておく方法が主に用いられる。特にWebサービス等のコンテンツ提供サービスにおいて、多数のユーザによって頻繁に利用されるデータは、オリジンサーバよりもユーザに近い位置に置かれたキャッシュサーバに記憶しておき、オリジンサーバに代わってキャッシュサーバからユーザへとデータを配信する。これ

によりネットワークのコア部分におけるトラフィック量を低減し、ネットワークリソースの有効活用を図るとともに、ユーザへの応答時間を短縮し、サービスの可用性を向上させることができる。

サービスの多様化に伴い、ユーザ情報やサービスによるトラフィックの識別が必要不可欠となり、現在ではアプリケーション層の情報を利用してトラフィックを振り分ける高位レイヤ対応のスイッチ（レイヤ7スイッチ）が広く利用されている。これらはサーバの負荷を低減する為に用いられる。それぞれ対応するサービスが異なるサーバを複数台用意しておき、ユーザからのリクエストの内容あるいはサーバの負荷状況に応じてトラフィックを振り分けることで、個々のサーバのパフォーマンスを向上させることができる。

また、同一のコンテンツを地域的に離れた場所に備えておき、データの配信先となるユーザのネットワークアドレスやサーバの負荷、応答時間に基づいて最適なサーバを選択し、当該サイトへリクエストを振り向ける手段も用いられる。なお、サーバの負荷分散技術はサイト内での局所的なトラフィック制御技術であり、ルーティングのようにネットワークアドレスにもとづく経路制御とは異なる。現段階ではトラフィックを低減するための技術に注目が集まっており、データ転送時のネットワーク制御に関する技術は存在しない。

【 0 0 0 3 】

【発明が解決しようとする課題】

近年のアクセス回線のブロードバンド化とストレージ装置の記憶容量の飛躍的な増大、さらには移動通信端末の普及により、ネットワークを通じての動画や音声等マルチメディアデータのトラフィックが増え、業務データの異動も頻発している。ネットワーク帯域の拡大に伴い、大容量データがネットワーク経由で利用されるようになると、網内各所で頻繁にバースト的なトラフィックが発生することが予想される。またストリーミングデータ等、一定時間ネットワーク帯域を流れつづける形のトラフィックも増大する傾向にある。

従来のネットワーク層におけるルーティングでは、ネットワーク上の経路がルーティングテーブルの中で一意に決定されることから、トラフィックは特定の経路に集中し、回線負荷が偏る傾向がある。特に複数のトラフィックが同時にネット

ワーク上のある区間に集中した場合に、宛先検索処理およびスイッチング処理の性能上の制約により、輻輳あるいはそれに伴う回線障害が発生する可能性がある。ルータは大容量データを低遅延、低損失率で配信する必要があるが、上記の理由により、従来のネットワークシステムでは、大容量データ配信に伴うトラフィックの増大に対して十分に対応することが困難である。

大容量コンテンツ配信に対応した技術であるCDNでは、サーバの分散およびキャッシュの分散配置によってリクエストに対するレスポンスを向上させるが、データ配信時には上記ネットワーク層ルーティングプロトコルを利用している為、経路制御に関して同様の課題が残っている。すなわち、同時に複数のトラフィックが発生した場合にトラフィックが集中し、転送効率が低下する。これに対し、トラフィックエンジニアリングによる経路制御をCDNと組み合わせるIP-VPNサービスの需要が拡大しており、帯域制御への関心が高まっている。しかし、トラフィックエンジニアリングでは経路設定後の帯域利用率の変化に対応する機能が準備されていないことから、トラフィックが重なった場合に処理が特定の回線に集中する問題は解決されずに残っている。ダウンロード中心の利用形態から、ストリーミングデータおよびリアルタイムコミュニケーションが主流となる将来のネットワークにおいて、解決しなければならない重要課題の一つである。大容量データを配信するために帯域を効率的に利用できる機能がネットワーク装置には必要である。

以上の点から、本発明では大容量データの配信に際してパケットロスおよびそれに伴う配信遅延を低減すること、ネットワーク内部での輻輳を低減することでネットワーク全体の配信効率を向上させること、および現状のネットワークに対して容易にルータの追加、削除ができ、容易にネットワークの性能向上を図ることが可能な情報ルーティング装置および方法を提案することを目的とする。

また、本発明は従来のネットワーク技術と連携することにより、コンテンツの属性に応じたQoS制御、ネットワークの利用状況に応じた柔軟な経路制御手段を提供し、データ配信効率を向上させることを目的とする。

【 0 0 0 4 】

【課題を解決するための手段】

本発明による情報中継装置は、複数の入力ポートを持つ受信部と、複数の出力ポートを持つ送信部と、データ処理部と、前記受信部と送信部と前記データ処理部を接続するスイッチ部と、記憶機能を有する複数のストレージ部と、前記受信部、前記送信部、前記スイッチ部、前記ストレージ部を制御するルーティン処理部と、から構成されるアプリケーション層でルーティング機能を持つ情報中継装置であって、

ネットワーク上で前記複数ストレージ部を識別するためのストレージアドレスと、前記ストレージアドレスと、データ属性と、を含むデータパケットを用いる情報中継手段を特徴とし、

前記受信部は、データパケットを受信する手段と、前記データパケットからストレージアドレスと前記データ属性を取り出す手段と、前記データ属性をルーティング処理部に転送する手段と、前記データパケットを前記スイッチ部に送り込む手段とを有し、

前記ルーティング処理部は、ストレージアドレスを含む経路情報に基づいて受信データパケットの送出先を、前記送信部または前記データ処理部から選択し、前記スイッチ部に対して切替を命令する手段と、

前記ストレージ部は、前記受信データを記憶する手段とを有し、

前記スイッチ部は、前記制御信号に基づき経路の切替を行う手段を有し、

前記データ処理部は、データ属性に基づいて、データを格納もしくは送信する手段を有し、

前記送信部は、前記制御部からの制御信号に従いパケットのヘッダ処理を行う手段と、データパケットを隣接中継装置へ転送する手段と、

を備えることを特徴とする。

そして、前記情報中継装置は、ネットワーク上でストレージ部を識別する為の識別情報（ネットワークストレージアドレス：NSA）を用いて表される経路表（ストレージルーティングテーブル：SRT）を制御部に保持する手段と、SRTを用いて当該宛先NSAに対応する送出先NSAを決定する手段と、を有することを特徴とする。

また、前記NSAは、中継装置のネットワーク上における位置情報と、物理的ある

いは論理的に区別される記憶装置の個体識別情報と、保存領域内においてディレクトリまたはブロックアドレス指定によるデータ保存位置を識別する情報と、のうち一つまたは複数の組み合わせによって表されることを特徴とする。

また、前記情報中継装置は、ストレージ部にメモリあるいはHDDの一つ又は複数の媒体から構成される記憶領域を有することを特徴とする。

また、前期情報中継装置は、スイッチ部において受信データをストレージ部に保存する為、入力データパケットをデータ処理部に送る手段と、前期ストレージ部は前期データ処理部からデータを受け取り、少なくとも次段の中継装置への転送が完了するまで受信データを装置内に保存しておく手段を有することを特徴とする。

本発明による情報ルーティング方式は、複数の入力ポートを持つ受信部と、複数の出力ポートを持つ送信部と、データ処理部と、前記受信部と送信部と前記データ処理部を接続するスイッチ部と、記憶機能を有する複数のストレージ部と、前記受信部、前記送信部、前記スイッチ部、前記ストレージ部を制御するルーティング処理部と、から構成されるアプリケーション層のデータ属性をもとにルーティングを行う情報ルーティング方式であって、

ネットワーク上で前記複数ストレージ部を識別するためのストレージアドレスと、前記ストレージアドレスと、データ属性と、を含むデータパケットを用いる情報ルーティング方式を特徴とし、

前記受信部は、データパケットを受信するステップと、前記データパケットからストレージアドレスと前記データ属性を取り出すステップと、前記データ属性をルーティング処理部に転送するステップと、前記データパケットを前記スイッチ部に送り込むステップとを有し、

前記ルーティング処理部は、ストレージアドレスを含む経路情報に基づいて受信データパケットの送出先を、前記送信部または前記データ処理部から選択し、前記スイッチ部に対して切替を命令するステップと、

前記ストレージ部は、前記受信データを記憶するステップとを有し、

前記スイッチ部は、前記制御信号に基づき経路の切替を行うステップを有し、

前記データ処理部は、データ属性に基づいて、データを格納もしくは送信するス

テップを有し、

前記送信部は、前記制御部からの制御信号に従いパケットのヘッダ処理を行うステップと、データパケットを隣接中継装置へ転送するステップと、
を備えることを特徴とする。

そして、前記制御ステップは、ネットワーク上でストレージ部を識別する為の識別情報（NSA）を用いて表される経路表（SRT）を制御部に保持する手段と、SRTを用いて当該宛先NSAに対応する送出先NSAを決定する手段と、を有することを特徴とする。

また、前記NSAは、中継装置のネットワーク上における位置情報と、物理的あるいは論理的に区別される記憶装置の個体識別情報と、保存領域内においてディレクトリまたはブロックアドレス指定によるデータ保存位置を識別する情報と、のうち一つまたは複数の組み合わせによって表されることを特徴とする。

また、前期制御ステップは、受信部において入力パケットのヘッダ処理を行うと共に受信データを制御部内メモリまたは記憶装置内に保存する為、入力データをデータ処理部に送る手段と、データ処理部およびストレージ部はヘッダ情報をもとにデータパケットを識別し、少なくとも次段のNSAへの転送が完了するまで受信したデータパケットの情報を装置内に保存しておく手段を有することを特徴とする。

【 0 0 0 5 】

【発明の実施の形態】

本発明の実施形態について詳細に説明する。図1は本発明を適用する装置の基本構成を表すブロック図である。

本発明における情報中継装置は、大容量データトラフィックが頻発するネットワークコアに利用されることを前提とする。装置間を接続し経路制御を行う為、複数の回線4-1～4-nに接続された複数の入力ポート（論理ポート）2-1～2-nを含むデータ受信部2と複数の回線5-1～5-nに接続された複数の出力ポート3-1～3-nを持つデータ送信部3とを有し、入出力の経路を装置内部で切り替える為、入出力ポートを関連付けるソフトウェアスイッチ部1を有する。入力および出力用の論理ポート2-1～2-n、3-1～3-nを通じてト

ランスポート層以下のプログラムと通信し、データの送信を行う。

スイッチ部はルーティング処理部での経路検索結果に基づいてデータパケットの送信に必要な入力ポートと出力ポートの組み合わせを決定し、装置内転送経路を生成する機能を有する。スイッチ部はまた、データの中継する際に一時的にその内容を記憶しておくことを目的として、制御部へ受信データの複製を送信する機能を備える。

ルーティング処理部 1 0 にはデータパケットに含まれるアプリケーション層のデータ識別情報にもとづきデータパケットの送信先を決定する経路検索処理を行う機能と、検索結果にもとづいてスイッチ部および送信部に対して制御信号を送信する機能を有する。ルーティング処理部 1 0 にはルーティングの為の経路表 SRT 1 1 を保持する。

データ処理部 3 0 は送受信時のデータの組み立ておよび分割、受信データの一時記憶機能を有する。ストレージ部へのデータの読み書きはデータ処理部を通じて行う。データの組み立ておよび分割処理に必要な情報はデータパケットのヘッダ部に含まれ、受信部および送信部にてヘッダ情報の取り出しあるいは書き込みが行われる。

データ処理部 3 0 に接続される記憶装置 2 0 は、大容量データを転送する際にデータを保管し、あるいは転送処理の際に主記憶領域が不足した場合にデータを蓄積しておく為の一つまたは複数のストレージ 2 1 で構成される。ストレージ 2 1 は大容量のメモリであってもよく、ハードディスクで構成しても、その組み合わせでもよい。以降の説明ではこれらの構成を含めてストレージと称することとする。

次に、本発明の装置内部における信号の処理について説明する。入力ポート 2 - 1 ~ 2 - n に供給されたデータパケットは受信部 2、スイッチ部 1、送信部 3 の順に通過し、出力ポート 3 - 1 ~ 3 - n を通じて次の中継装置へと転送される。受信部 2 は受信したデータパケットに含まれるヘッダ情報を抽出し、それらを制御情報としてルーティング処理部 1 0 に送る。ルーティング処理部 1 0 ではストレージルーティングテーブル SRT 1 1 によって受信部 2 から受け取った宛先装置アドレス等の制御情報をもとに送信部 3 内の出力先ポートを決定し、スイッチ部

1 および送信部 3 へ制御信号を送る。制御信号によってスイッチ部 1 は入力ポートと出力ポートの対応を決定し、データパケットは送信部 3 へと送られる。受信部 2 において抽出される制御情報にはパケットの宛先装置アドレス、送信元装置アドレスおよびその他のデータ属性を示す識別子が含まれる。送信部 3 ではルーティング処理部 1 0 より受け取った制御情報にもとづいて必要ならばヘッダ情報を付加し、次の中継装置へ送出する。

図 1 に示された構成は主にソフトウェアにより構成され、ISO 参照モデルにおけるトランスポート層よりも上位の層における処理のみで記述される。経路決定は上位層で行われ、転送に際しては受信部および送信部を通じて下位層のプロトコルと通信を行うことにより実行される。

また、図 1 の構成は受信部および送信部に物理インタフェースを用い、スイッチ部に、例えば $N \times N$ のクロスバースイッチを用いた情報中継装置の設計も可能である。この場合は転送に関わる受信部、スイッチ部、送信部はネットワーク層アドレスを用いる従来のルータアーキテクチャと共通の構成となる。受信部および送信部にはパケット処理を専用に行う為に例えば高速ハードウェア処理等を行う回路が組み込まれ、SRT を用いたアプリケーション層での経路決定とデータリンク層以下の物理アドレスとが NSA と直接連携する方式での制御が行われる。

図 2 は本発明に従った、情報ルーティング処理のフローチャートである。

データパケットは通信回線 4-1 ~ 4-n から入力ポート 2-1 ~ 2-n を通じて受信部 2 へと供給される。ステップ 1 0 1 において受信部 2 がパケットを受信すると、ステップ 1 0 2 でパケット上のアプリケーション層データが収められる領域の先頭部分に書き込まれた宛先情報等の制御情報（以下、データヘッダと称する）が読み出される。ヘッダに含まれる宛先情報と自装置の持つアドレスとを比較し、ステップ 1 0 3 の判断によりデータの宛先が自装置のアドレスと一致した場合には、ステップ 1 0 4 においてデータヘッダを含むヘッダ部分の情報に基づき、必要ならば複数のパケットで送られてくるデータを結合して、受信したデータの種別に応じ自装置内に保存あるいは実行する等のデータを運用する為の処理を行う。ステップ 1 0 3 にて受信パケットが自装置宛てではなく、自装置を経由して宛先に向かうデータパケットであると判断した場合、ステップ 1 0 5 によってルーティ

ング処理部 1 0 に保持されている SRT 1 1 により宛先アドレスに対応する次の中継装置の NSA を決定し、パケットの入力ポートと SRT 1 1 で得られる出力ポートの組み合わせをスイッチ部 1 へ送信し、この制御信号によってパケット送信に適するスイッチ部の入出力ポートの接続が確立され、データが送信部へと送られる。次いでステップ 1 0 7 では送信部 3 において、ルーティング処理部 1 0 で決定されたパケットの送信経路に従ってデータヘッダが付加され、次の中継装置へ送信される。送信後、ステップ 1 0 7 によってデータ転送の成否を判定し、次の中継装置と自装置との間で何らかの異常によりデータの一部または全部の送信に失敗したと判断された場合にはステップ 1 0 8 により装置内部に蓄積されているデータから再送処理を行う。正常にデータ転送が終了したと判断された場合にルーティングのプロセスは終了する。

図 3 は本発明の情報ルーティング方式に関するプロトコルの相互関係を示した概念図である。ストレージアドレスのレイヤで情報中継装置を結んだネットワークを形成する。隣接装置間での自律転送式のネットワークという形態であるが、実施例においては転送の状態を把握し、さらに確実な転送を行う為、送信元と宛先との間にセッションを確立しておくことが考えられる。自装置のストレージ部 2 0 から次の中継装置内の記憶装置への転送にはトランスポート層プロトコル（TCP 等）の高信頼転送プロトコルを利用することで、中継装置間の転送を保証することができる。あるいは本発明の実施例に示されるようにそれぞれの情報中継装置においてデータをストレージ部に格納する際にデータの完全性を確認する手段を備えておいてもよい。隣接中継装置間で何らかのデータ完全性の保証を提供する機能を有するものとする。従ってある区間でのデータ送信が終わる毎に装置間のネットワーク帯域は開放される。

図 4 に本発明の情報ルーティング方式で利用するデータヘッダの基本要素を図示する。データパケットはデータヘッダ部 3 0 1、データ部 3 0 2 から構成される。データヘッダ 3 0 3 には少なくともストレージネットワーク内での最終的な宛先となる情報中継装置のネットワークストレージアドレス（NSA）3 0 3 および送信元となる NSA 3 0 4 が含まれる。ヘッダ情報 3 0 1 は宛先アドレス検索、送信時のデータ分割、受信時のデータ再生等の処理に利用される。

図 5 に本発明に従った情報ルーティング方式におけるデータの分割処理を示す。ステップ 1 2 1 でストレージ部 2 0 から読み出したデータはステップ 1 2 2 において送信に適当な一定の大きさを分割され、ステップ 1 2 3 および 1 2 4 にて分割されたそれぞれのデータ片に必要なヘッダを付加して送り出される。データは情報中継装置を経由するたびに各装置のデータ処理部 3 0 においてデータの再生あるいは分割、転送が行われるが、各中継装置においてデータ受信から送信に至る一連の処理が終了し次第、記憶領域は解放される。

図 6 は本発明の情報ルーティング方式で利用するストレージルーティングテーブル (SRT) の基本構成を示す図である。SRT はルーティング処理部 1 0 のメモリ内に蓄積され、受信部 2 からルーティング処理部 1 0 へ送られるデータパケットの宛先 NSA 情報 3 0 1 に対してテーブル上で対応するアドレス 4 0 3 を検索し、適した送出先アドレス 4 0 4 を決定する際に用いられる。SRT から得られた送出先アドレスは該当するパケットの入力ポート番号と対応づけられ、スイッチ部 1 の制御に用いられる。

SRT は、少なくとも本発明によるルーティング方式で転送される経路上の最終目的アドレス 4 0 1、そこに向かう為に通過するべき次の中継装置のストレージアドレス 4 0 2 を含む。さらに、参照情報として経路上に存在する中継装置内の記憶領域の空き容量、経路更新からの経過時間等の補助情報が含まれていてもよい。

本発明によるルーティング方式では、中継されるべきノードの記憶領域の空き容量に基づき、各ノードの記憶領域を最も有効に利用できる経路を選択する。本方式では IP 層のルーティングテーブルとは独立に経路を決定することが可能となる。テーブルはネットワークトポロジや中継装置の記憶領域の利用状況に対応できるように、定期的に更新されることとするが、随時必要な時に更新を行うことも出来る。

図 7 は受信データから中継装置内の記憶領域にデータを保存する処理における処理の流れを示す。図 2 に従い、受信部 2 に供給されたデータパケットはコンテンツヘッダ処理 1 0 2 において、宛先、属性情報の確認が行われる。データを保存するかどうかの判断 1 4 1 は、情報ネットワークでの中継記憶装置の空き領域等

の外部要因と、中継されているデータの持つ属性等の個々のヘッダ情報とによって行われる。ストアしないと判断した場合には、ルーティング処理のみを行う。データを保存する場合、ステップ 1 4 2 でコンテンツヘッダに含まれるコンテンツ識別子およびデータ分割・送信時のパケット通し番号等の再生用情報によってデータを再構築する。この時、データ転送を確認するステップ 1 4 3 によりデータの一部分が受信されていないと判断される場合にはステップ 1 4 5 の再送要求を前段あるいは最も近い位置にある中継装置へ送信し、ステップ 1 4 6 で損失分のデータを取得する。再生されたデータがステップ 1 4 4 においてルーティング処理部 1 0 あるいはストレージ部 2 0 に保存されると一連の処理は終了する。

図 8 はデータ再生、保存処理に係るデータパケットのヘッダ情報を図示したものである。情報ルーティング時のアドレス情報 3 0 3 および 3 0 4 に加え、データ識別子 3 0 5、通し番号等のデータ分割時の識別子 3 0 6 を含む。この場合、通し番号 3 0 6 はデータ識別子 3 0 5 に付随する。図 7 のステップ 1 4 2 におけるデータ再生時の処理を図 9 に示す。データはステップ 1 6 3 でデータ識別子 3 0 5 によって区別され、同じ識別子をもつパケットはパケット識別情報 3 0 6 によって正しい並び順に整理され、ステップ 1 6 4 においてデータが再生される。

図 1 0 に再生時のデータパケット確認処理を示す。データ再生時に組み立て情報として用いる分割番号 3 0 6 により、必要なパケットを確実に受け取っているかどうかをステップ 1 8 3 で判断し、データパケット不着、または不完全なパケットで受信された場合には前段の中継装置へ再送要求 1 8 5 を行う。

図 1 1 に中継装置が送出先からデータ再送要求を受け取った場合の処理の流れを示す。ステップ 2 0 1 で送信要求を受信すると、通常のルーティングと同様にコンテンツヘッダの処理 2 0 2 を行う。次いで 2 0 3 において要求されているデータと要求される分割番号等の情報に基づいて該当するデータをメモリあるいはストレージから読み出し、データパケット作成の為に分割した後、宛先 NSA 情報を含むデータヘッダの処理 2 0 4 を経てステップ 2 0 5 で次の中継装置へ送信する。送信の成否を判定するステップ 2 0 6 で送信完了と判断されると送信処理は終了する。

図 1 2 に送信要求を伝えるパケットのヘッダ構造を示す。このヘッダには要求するデータを識別する為の識別子 3 0 5 と必要とするデータの分割番号 3 0 6、およびこのパケットが送信要求信号を伝えるものであることを示す信号識別子 3 0 7 を含む。

図 1 3 に本発明の実施例として NSA の定義例を示す。ネットワークストレージアドレス NSA は物理的に区別可能な HDD 等のストレージ装置の識別情報としての利用に加え、パーティションで論理的に区別される領域としての定義、さらにはデータの格納場所を示すディレクトリ指定を用いた定義も含む。従ってインタフェース（ポート）と NSA の対応は NSA の定義により多様に変化し、柔軟なネットワークを構築することが可能である。

図 1 4 は各中継装置が複数の NSA を持つ場合のネットワーク構成例を説明する概念図である。情報ルーティング網は各中継装置の持つ NSA を用いて表される為、例えばある中継装置が複数の NSA を持つ場合はデータ中継用領域が確保できる NSA を必要に応じて選択することにより、複数のデータフローを同時に中継することができる。図 1 4 ではサーバ 5 1 3 からエッジ装置 5 0 5 を介してユーザ 5 1 1、5 1 2 にデータ配信する状況を仮定している。中継装置 5 0 1 は 2 つの NSA を持ち、データフロー 5 1 4 a 用に NSA が「a」で表される領域を使用し、装置 5 0 5 では NSA が「c」の領域を使用している。フロー 5 1 4 a の終了を待たずにデータフロー 5 1 4 b が発生した場合、中継装置 5 0 1 では領域「b」を使用することで転送用領域を確保し、エッジ 5 0 5 では例えば領域「e」を使用してユーザまでの経路を確立する。

本発明に従う情報中継装置では情報ルーティング機能と同時にデータを記憶装置に保存する。情報を保存しつつ転送することにより、従来のネットワーク層プロトコルによる転送方式と比べ、より確実なデータ配信が可能である。

図 1 5 に受信データのデータ処理部への保存とルーティングに関する処理のフローチャートを示す。これは、受信データサイズが小さい場合もしくはストレージに格納する必要のないデータのフローを受信した場合の処理に相当する。受信部 2 によって宛先確認およびデータパケット識別情報の抽出等、データヘッダ 3 0 1 の処理が行われると、ヘッダ情報に基づいた一時保存処理 6 0 3 が行われ、再

生されたデータはデータ処理部 3 0 に保存される。以下、ステップ 6 0 4 ~ 6 0 7 において図 2 に示されるフローチャートによってルーティング処理が行われ、保存されているデータ処理部 3 0 内のデータは、送定の完了が確認された段階で消去される。再送処理 6 0 6 では、データ処理部に保存されているデータパケットもしくは組み立てられたデータを読み出し、必要ならば分割およびヘッダ付与が行われた後に送定される。

図 1 6 は受信データをストレージ装置に保存する処理を示す。ステップ 6 2 5 に表されるデータ保存およびルーティングの処理は図 1 5 に説明される通りである。データをデータ処理部 3 0 で構築し、ステップ 6 2 9 において自装置のストレージ部 2 0 に保存領域が十分にあるかどうかを判断する。空き領域が少ないと判断される場合にはステップ 6 3 0 で一時データ処理部 3 0 内のメモリ上に保持しておく。ストレージへの保存が必要なデータは図 1 5 の場合にメモリ上で一時保管されるデータとは異なる種類のものと考えられる。ストレージ部 2 0 の記憶領域が空き次第、ストレージ部 2 0 への保存および当該装置からのデータ配信などの運用処理 6 3 1 が行われ、この処理は終了する。なお、データパケットが自装置宛ではない場合には図 1 5 の場合と同様にルーティング処理後メモリ上のデータ領域は解放することができる。

図 1 7 では、データ処理部 3 0 内のメモリとストレージ部 2 0 とを連携してデータ保存に利用する場合の処理を示す。ステップ 6 4 2 でメモリ空間にデータ再生に十分な領域がない場合はステップ 6 4 3 でストレージ領域を確認した後、メモリ上で保存されているデータをストレージ部 2 0 内に移動し、データ処理部 3 0 内の空き領域でデータを構築する。データはステップ 6 4 7 において再生され、メモリあるいはストレージ内で保存されるか、運用処理が行われる。メモリ内のデータを保存する為に十分なストレージ領域がない場合、受信パケットが自装置宛か否かをステップ 6 4 5 で判断し、自装置宛であれば廃棄し、再送処理を要求する。自装置宛でなければ、保存処理を伴わずにルーティング処理 6 4 8 を行うことによって、受信データを次の中継装置へ送信する。ルーティング処理 6 4 8 に続いて行われる処理は図 2 に従う。

本発明の情報ルーティング方式を適用することにより、アプリケーション層の情

報に基づく柔軟な経路変更を行うことができる。

図 1 8 は本発明の実施例の一つとして、データヘッダに含まれるコンテンツ識別子を用いた、優先レベルの異なるデータの転送を行う場合の経路制御のフローチャートである。入力ポートに供給されたデータパケットに対し、ステップ 6 6 1 にて受信パケットから抽出されたコンテンツ識別子を用いてステップ 6 6 2 で優先度レベルを決定し、ステップ 6 6 3 で優先度および必要帯域に基づく経路選択処理によって適切な経路へと送出する。判断 6 6 4 において当該データを転送する上で確保可能な帯域が十分でないと判断された場合にはステップ 6 6 5 により別経路を設定し転送する。

図 1 9 に優先度別経路制御を実現する為のデータパケットのヘッダ構成例を示す。データヘッダ部 3 0 1 には宛先 NSA 3 0 3、送信元 NSA 3 0 4、データ識別子 3 0 5 に加えて該データの優先度識別情報 3 0 8 を含む。

経路情報テーブルとコンテンツ識別子との連携方法は、例えば情報ルーティングテーブル内にコンテンツ識別子を含むものを用意してもよく、優先度別に複数の SRT を用意しておく方法でもよい。

図 2 0 にデータパケット優先制御の為の SRT の構成例を示す。

本発明における情報ルーティング方式ではデータを各中継装置に保存しつつ転送するため、回線障害時に送信元までデータの再送を要求する必要がある。

図 2 1 を用いてこの Hop-by-hop でのデータ再送機構について説明する。ここでは、本発明における情報中継装置 A 7 0 1、B 7 0 2、C 7 0 3 がこの順に接続されていると仮定する。通常の情報ルーティングにおいては送信 7 0 4、7 0 5 によってデータは各中継装置のストレージ部に保存されつつ次の中継装置へと転送される。以下、再送の生じる例として二つの場合を取り上げる。データがある区間で完全に失われる場合 (Case 1)、中継装置 C 7 0 3 はデータフローの存在を知り得ない。中継装置 B 7 0 2 はデータ送信の成否を確認する為に一定時間待ち、確認応答がなければ再度データを送信する。この時必要があればデータの中継経路の再設定を行う。データの一部が失われた場合 (Case 2) はデータ受信の中継ノード C 7 0 3 においてデータの再生、およびそれに伴うデータの完全性確認処理が行われる。中継装置 C 7 0 3 は損失パケットのもつデータヘッダ情報を確認し

、前段階の中継装置B702に再送要求712を送り、B702にデータが保持されていれば直ちにC703に向かって再送713を行う。装置B702にデータが保持されていない場合にはさらに前段の装置A701に再送要求714が送られ、必要なデータを再送715、716によって中継装置Aから中継装置Cへ送信する。

図22によりHop-by-hop方式を用いた場合の効果を説明する。サーバ810から配信されるデータフロー813はエッジ装置801、コア中継装置802、エッジ装置805を介してユーザ端末807に到達する。同時にサーバ810からエッジ中継装置801、コア中継装置803、804、エッジ装置812を経由して端末809へのデータ送信812が行われているとする。エッジ装置806とコア装置804の間で回線障害によってデータの一部が失われた場合、エッジ装置806は再送要求をコア装置804に送り、装置804は自装置に保存されているデータから必要なパケットを送信する。従来のネットワークのように再送要求がサーバ810まで到達することがないため、サーバ810、エッジ装置801、コア装置802への処理の負荷が少なくなり、フロー813の転送効率に影響を及ぼさない処理が可能である。

図24のフローチャートを用いて、次に、本発明の実施例として、ネットワーク層の情報を経路決定に利用する場合の処理を説明する。

図23は本方式を適用する為のアーキテクチャの基本構成を示すものである。

受信したデータパケットの宛先に向かう経路は通常複数存在する為、可能な経路が複数検出された場合には何らかの手段を用いて送出先を決定する。例えばNSAに対して機械的に優先順位を付与する等の方法が考えられるが、ネットワーク層プロトコルの持つ障害、輻輳状況、および回線コスト等の情報を情報ルーティングでの経路選択に際して使用することでより効率的なルーティングが可能となる。

ステップ901で中継装置の入力ポートに受信されたパケットは、ステップ902において受信部でコンテンツヘッダ情報を抽出され、ステップ903で自装置宛でない判断されると、ステップ905で宛先情報に基づいて送信先が決定される。ここで送信可能な次の中継装置が複数存在した場合905にステップ90

7にて各中継装置への回線状況、あるいは装置内の処理にかかる負荷をICMP、SNMPといった下位層のプロトコルを利用して取得し、SRTと合わせて最も有効な経路を決定する。ネットワーク層の情報に関してはSRTに含んでおいてもよいし、SRTとは別に、ネットワークの状況を保持しておく為のテーブルを用意してもよい。

図25を用いてネットワーク層プロトコルと連携する際の情報中継装置の動作を説明する。サーバ926からユーザヘデータを転送するにあたり、エッジ中継装置921は次の中継装置922および923の持つストレージの空き容量を参照してSRTを作成する。中継装置921は複数の宛先候補を持つが、どちらかに空き容量が足りない場合には空いている方向へ転送し、どちらの装置もデータ転送に利用できる状況ではネットワーク層のプロトコルを利用することで、よりコストの少ない経路930aを選択する。さらにICMP等のネットワーク層における経路情報を参照し、例えば他のトラフィックによって経路930aの帯域が十分とれないと判断した場合にはSRTを経路930bへ切り替える。

本発明における情報中継装置は、既存ネットワークにおけるルータを全て置きかえる必要はなく、例えば一定のルータ数毎に設置してもよい。

図26はMPLSのトラフィックエンジニアリング(TE)機能を利用し、経路を決定する処理を説明するフローチャートである。本実施例においてはステップ941～945において、図2に従う処理によってデータパケットからヘッダ処理を行い、SRTを参照して転送先NSAを決定する。宛先が複数存在する場合もしくは次の中継装置に至る経路が複数あり得る場合、ステップ946で中継装置間の可能な経路について転送経路(Label Switched Path)の設定を行い、転送用経路を決定すると共に経路間の転送品質を保証する。

図27は本発明の情報ルーティング方式を現状の(IP)ネットワークに導入した際の情報中継装置内の処理を説明するフローチャートである。本発明による情報中継装置を、ネットワーク層でトラフィックが集中する個所に導入することにより、ネットワーク層プロトコルと共存することができ、ネットワークの拡張に際してコスト面のメリットがある。ステップ1001でパケットを受信した後、受信部2においてヘッダ情報を抽出する際、そのデータパケットが本発明に従う情

報ルーティング方式を適用可能か否かを判断するステップ1002を実行する。情報ルーティングを行うか否かの判断1002は、上位層ヘッダの有無、ネットワーク帯域利用状況、データ属性等により決定される。情報ルーティングを行わない場合、ステップ1003により既存のトランスポート層以下のプロトコルを用いて経路制御によって情報ルーティング網内を通過する。送信、再送処理のステップ1004はTCP、UDP等の従来の転送制御プロトコルによる。ステップ1002においてデータパケットが情報ルーティングを必要とすると判断された場合、ステップ1005～1009において図2に示す情報ルーティングの一連の処理が行われる。

図28は既存網に情報ルーティング方式を導入する場合の概念図を示したものである。エッジノード(EN)1101、1105は情報ルーティング網のエッジ装置であり、コアノード(CN)は情報ルーティング網のコア部分に置かれた中継装置である。サーバ1109より送信されたデータはエッジルータ1106を経由して網内に入り、図26に説明されるルーティング処理は情報ルーティング網内の各中継装置にて行われる。エッジ装置1101では上位ヘッダ、帯域情報、データ属性等を確認した上で情報ルーティング方式が必要と判断された場合には上位ヘッダの処理を行い、図2に従うルーティングを行う。上位ヘッダはサーバ1109によって付加される場合も、エッジ装置1101によって付加される場合もある。情報ルーティング網から既存網への転送に際しては対面エッジ装置1105によって必要な上位ヘッダ処理が施される。すなわちサーバと端末装置の間もしくは情報ルーティング装置1105と端末装置の間で情報ルーティングが行われる場合には上位ヘッダを保持しておき、情報ルーティング網内のみ上位ヘッダを使用する場合にはエッジ装置1105にて上位ヘッダは取り除かれる。

以上で説明した本発明の実施例によれば、ネットワーク上のデータ転送において、動画データ、音声データ、業務データ等の大容量データの転送を行う際に、ストレージの記憶領域を利用したデータ転送を行うことで、ネットワーク内のトラフィック状況に応じた適当な経路選択が可能であり、従来のルーティングにおいて経路が限定されることによる帯域不足とそれに伴う輻輳を回避することができる。また、NSAを導入し、ストレージの空き領域に基づくルーティングを行う

ことによりネットワーク層ルーティングに依存しない方式での経路設定が可能となり、従来と異なる帯域の確保ができる。このため、既存プロトコルとの共存が可能であり、それぞれの特徴を生かすことでより柔軟なネットワーク設計が可能である。また、アプリケーション層のデータ属性を利用するルーティングを行うことで、従来ルータのネットワークインタフェースに比較して細かいネットワーク構築ができる。さらに論理的に区別される複数のNSAを装置内に持つことができることから、ネットワーク層よりもアドレス付けの自由度が増し、データ転送において優先制御への応用が可能になる、アドレス管理が効率化される、といった効果が期待できる。

また、本発明では各情報中継装置にデータ蓄積機能を備え、ルーティングとデータ蓄積を同時に行なうことにより、大容量のトラフィックが発生する場合にも確実にデータを配信することができる。また中継装置間でデータフローの確認を行うため、大量のデータを転送中にネットワーク障害が発生し一部または全部のデータパケットに不着が生じた場合も最も近くの装置から必要なデータを再送することができ、従来のネットワークに大容量コンテンツが流れる際に頻発することが予想される多量の再送要求に伴うネットワーク資源の浪費を大幅に軽減することができる。

また、本発明は既存のネットワークに対して容易に装置を導入することが可能である。網全体で装置を入れ換える必要はなく、トラフィックの集中する場所に本発明を導入することにより帯域の有効活用ができ、コスト面での効果も期待できる。

【 0 0 0 6 】

【発明の効果】

従来のネットワーク技術では難しかった帯域の有効活用が可能となる。

【図面の簡単な説明】

【図 1】

本発明を適用する情報中継装置の基本構成を示すブロック図である。

【図 2】

本発明に従った、情報ルーティング方式に関する、データ転送処理のフローチ

ャートである。

【図 3】

本発明に従った情報ルーティング方式に関する通信プロトコルの構成を示すブロック図である。

【図 4】

本発明に従った情報ルーティング方式に関する、転送時に分割された個々のデータ片に付加されるヘッダの基本構成の図示である。

【図 5】

本発明に従った、情報ルーティングにおけるデータ分割およびデータパケット生成処理のフローチャートである。

【図 6】

本発明に従った情報ルーティング方式での経路制御に関する情報ルーティングテーブルの基本的な構成要素の図示である。

【図 7】

本発明の情報ルーティング方式に関する、受信データの再生、保存に関わる全体の処理の流れを示すフローチャートである。

【図 8】

本発明に従った、受信データの保存処理に利用するデータヘッダの基本構成の図示である。

【図 9】

本発明に従った、データパケットからデータを再生する処理のフローチャートである。

【図 1 0】

本発明の情報ルーティング方式においてデータの完全性をチェックする処理と再送要求を送信する処理のフローチャートである。

【図 1 1】

本発明に従った情報中継装置内に保存されたデータを送信する処理の流れを示すフローチャートである。

【図 1 2】

本発明に従った情報ルーティングにおいて、データ再送要求時に送信元情報中継装置に送信される再送信要求パケットのデータヘッダ構成の図示である。

【図 1 3】

本発明におけるストレージアドレスの定義例である。

【図 1 4】

本発明におけるストレージアドレスを利用したネットワーク構成の概念図である。

【図 1 5】

本発明の情報ルーティング方式に係る、ルーティングとデータ処理部における処理を示すフローチャートである。

【図 1 6】

本発明の情報ルーティング方式に係る、ルーティングとストレージ部へのデータ保存を行う場合の処理のフローチャートである。

【図 1 7】

本発明の情報ルーティング方式において、受信データを保存する際のデータ処理部とストレージ部との動作の連携を示すフローチャートである。

【図 1 8】

本発明の特徴に従って、コンテンツ別に優先度をつけて経路制御を行う際の処理を示すフローチャートである。

【図 1 9】

本発明の特徴に従って、経路の優先制御を行う際に用いられるデータヘッダの基本的な構成要素を示す図である。

【図 2 0】

本発明の特徴に従って、経路の優先制御を行う場合に中継装置によって参照されるストレージルーティングテーブルの構成例を図示したものである。

【図 2 1】

本発明の情報ルーティング方式における、隣接中継装置間でのデータパケットの再送処理の流れを表す図である。

【図 2 2】

本発明の情報ルーティング方式における、隣接中継装置間でのデータ再送処理の効果を説明するネットワーク図である。

【図 2 3】

本発明の一実施形態に従った、ネットワーク層ルーティングプロトコルと連携する転送経路制御を行う為の基本アーキテクチャを示すブロック図である。

【図 2 4】

本発明の一実施形態に従った、ネットワーク層経路情報を利用する情報ルーティング処理のフローチャートである。

【図 2 5】

本発明を用いたネットワーク層との連携による経路制御を行う場合の効果を説明する図である。

【図 2 6】

本発明の情報ルーティング方式とトラフィックエンジニアリングを組み合わせた処理の流れを示すフローチャートである。

【図 2 7】

本発明の情報中継装置を既存網に追加する場合の処理の流れを示すフローチャートである。

【図 2 8】

本発明の情報中継装置を既存網に追加する場合のネットワーク構成を示す図である。

【符号の説明】

1 スイッチ部

2 受信部

3 送信部

1 0 ルーティング処理部

1 1 ストレージルーティングテーブル

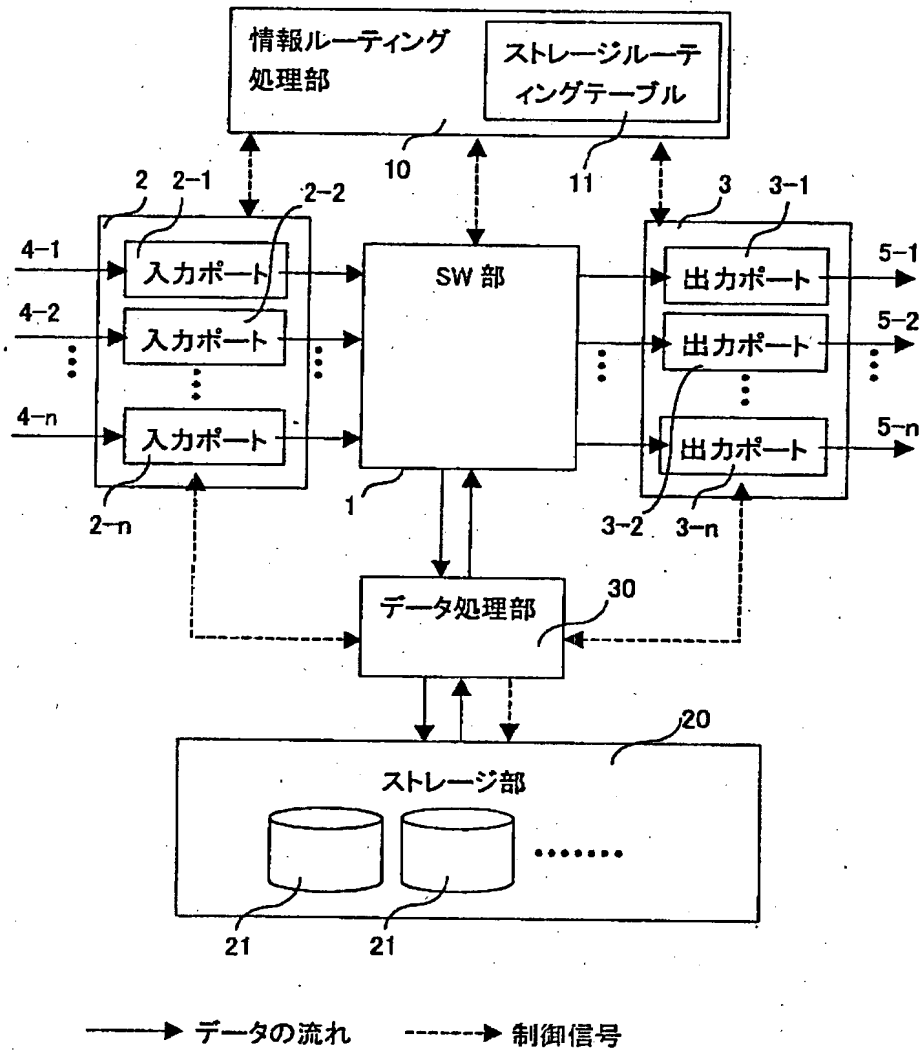
2 0 ストレージ部

3 0 データ処理部。

【書類名】 図面

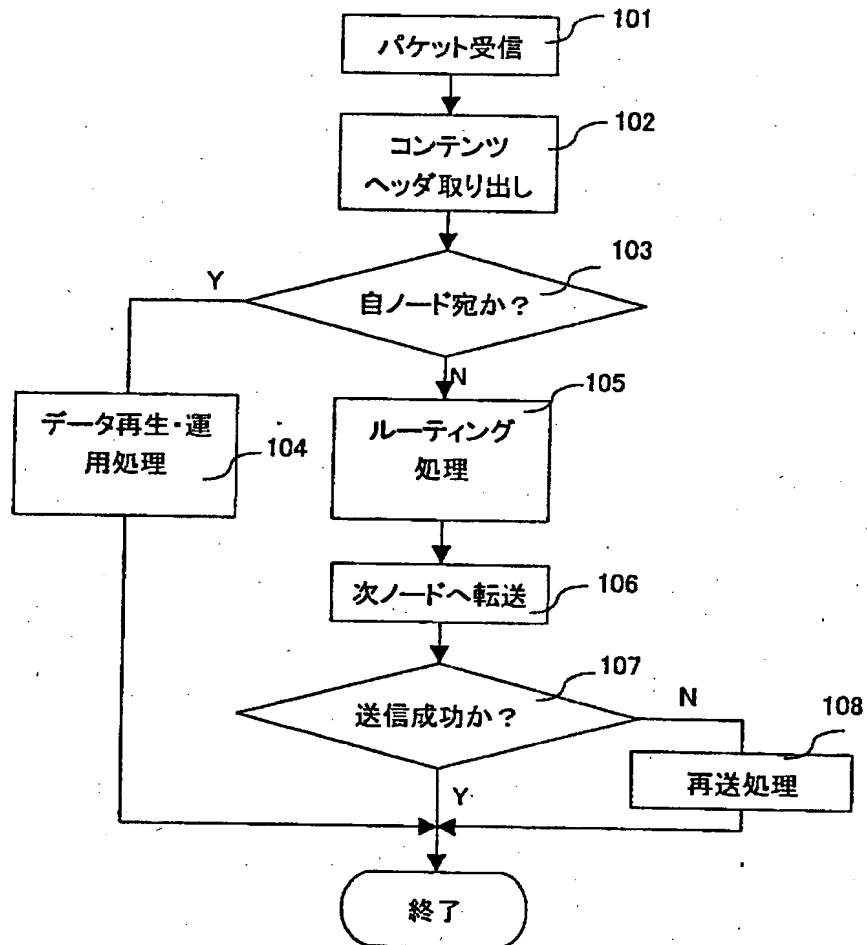
【図 1】

図 1



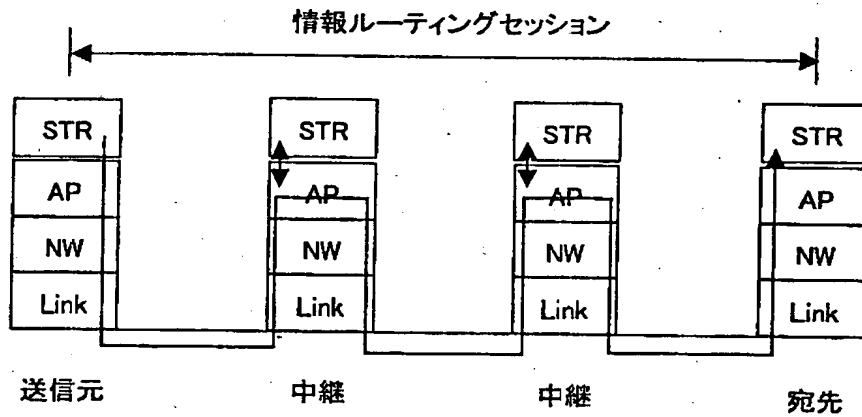
【図 2】

図 2



【図 3】

図 3



STR:ストレージ

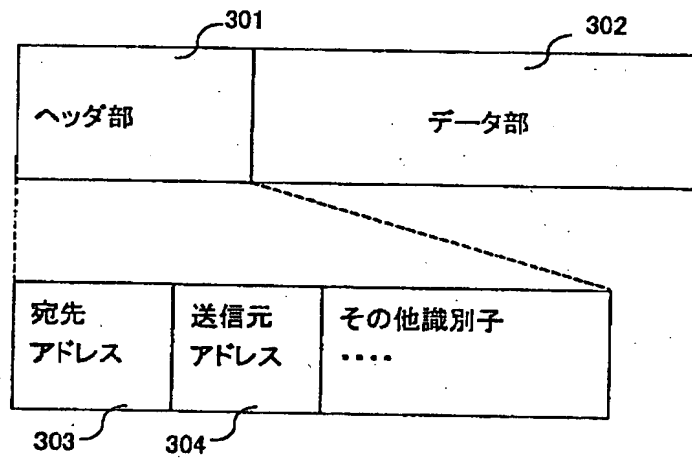
AP:セッション層以上の上位アプリケーション

NW:TCP/IP を含む転送用プロトコル

Link:DL 層以下のプロトコル

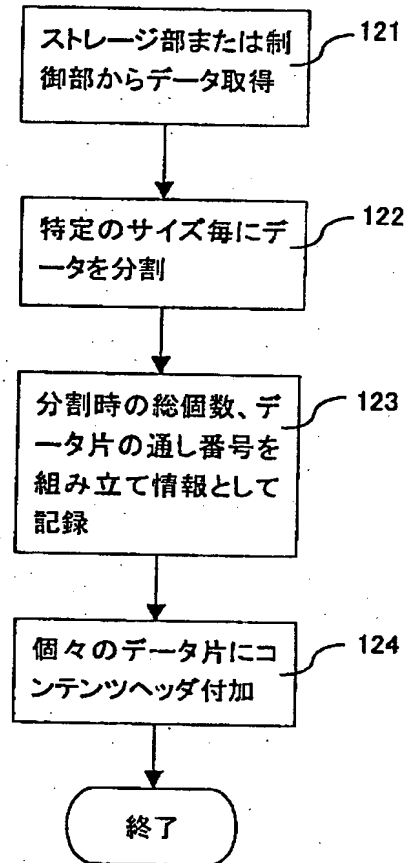
【図 4】

図 4



【図 5】

図 5



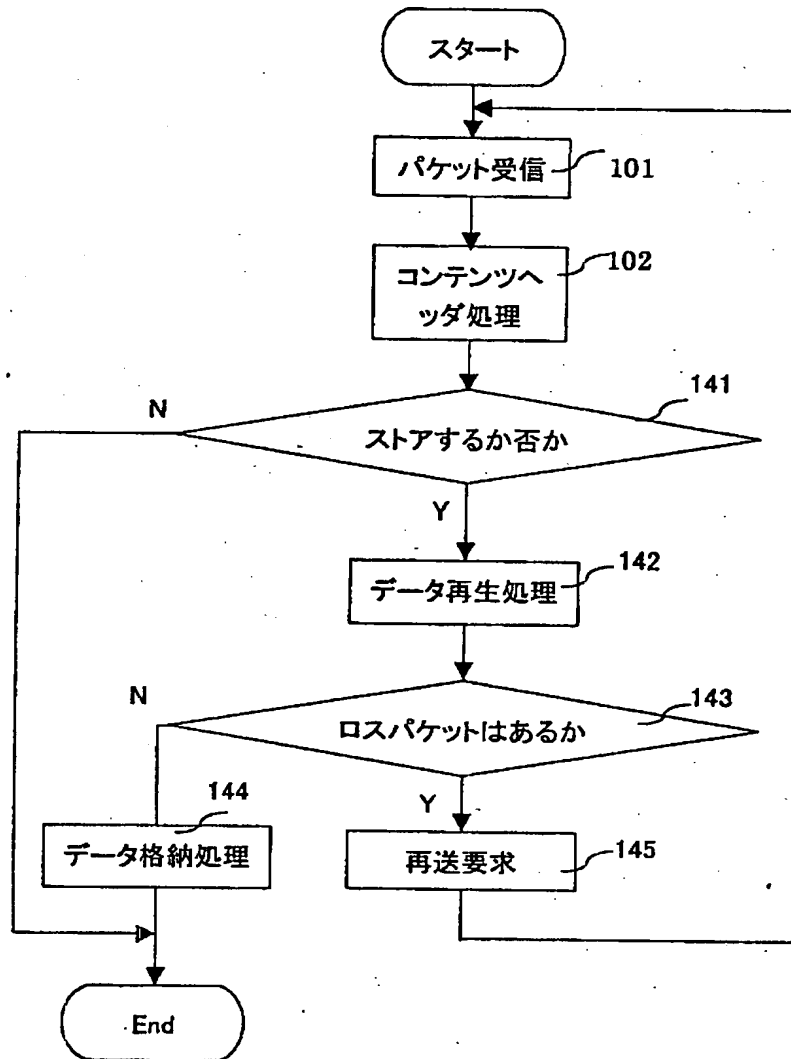
【図 6】

図 6

宛先アドレス	Next hop アドレス
A	a
B	b
⋮	⋮

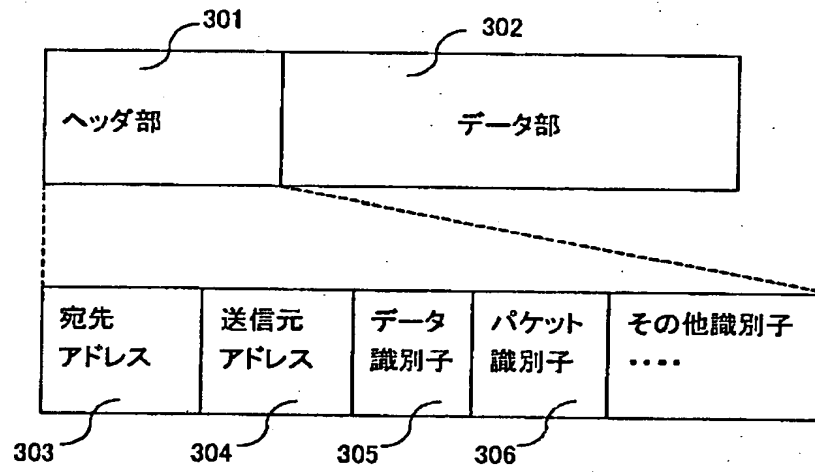
【図 7】

図 7



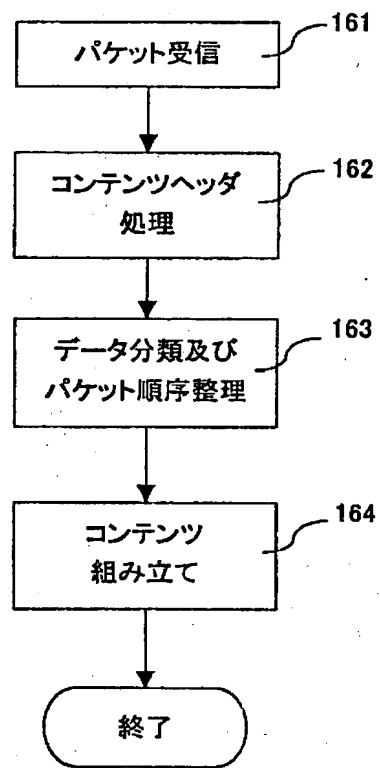
【図 8】

図 8



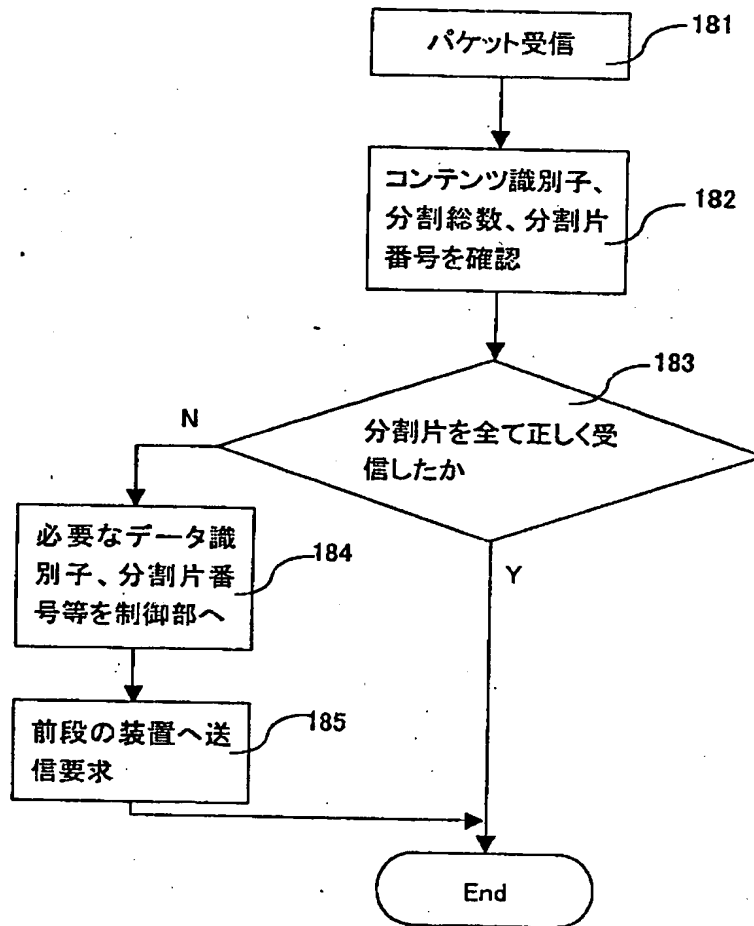
【図 9】

図 9



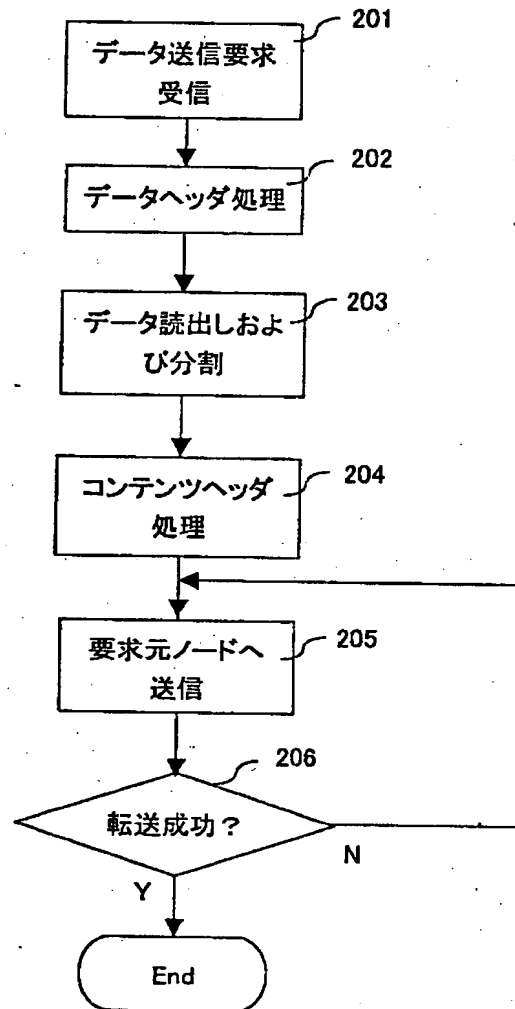
【図10】

図 10



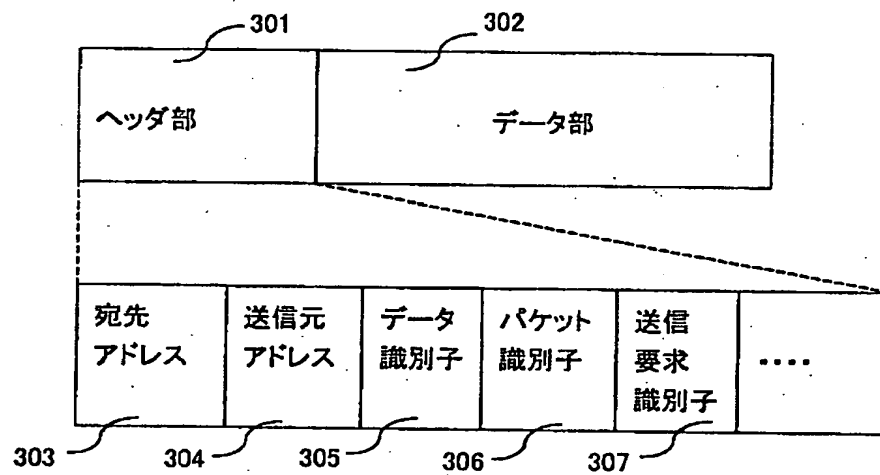
【図 11】

図 11



【図 12】

図 12



【図 13】

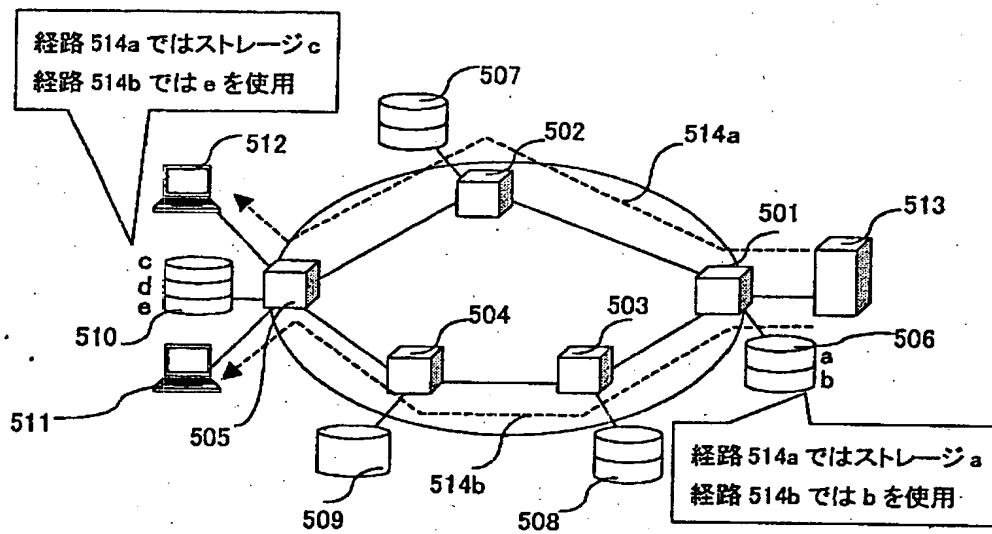
図 13

宛先アドレス	Next hop アドレス	格納ポイント
A	a	/home
B	b	0xbfbfd9e0
⋮	⋮	⋮

Diagram illustrating a table structure. The table has three columns: 宛先アドレス (Destination Address), Next hop アドレス (Next hop Address), and 格納ポイント (Storage Point). The rows are labeled with reference numerals 401, 410, 402, 411, and 403. The data rows are labeled with reference numerals 412 and 415. The vertical ellipsis (⋮) is labeled with reference numeral 413. The horizontal ellipsis (⋮) is labeled with reference numeral 414.

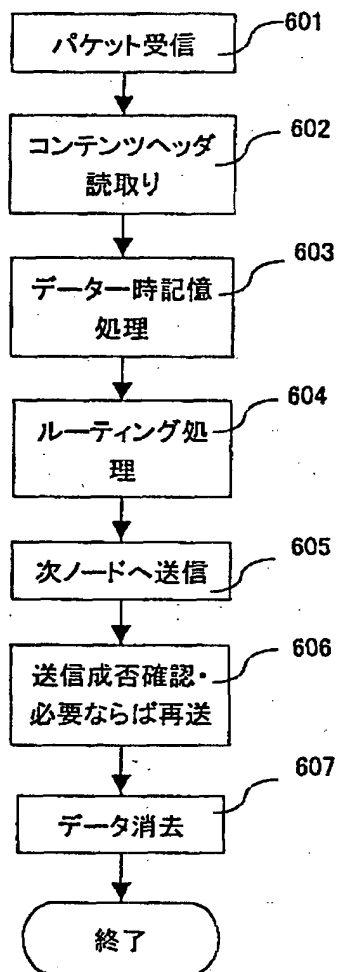
【図 1.4】

図 14



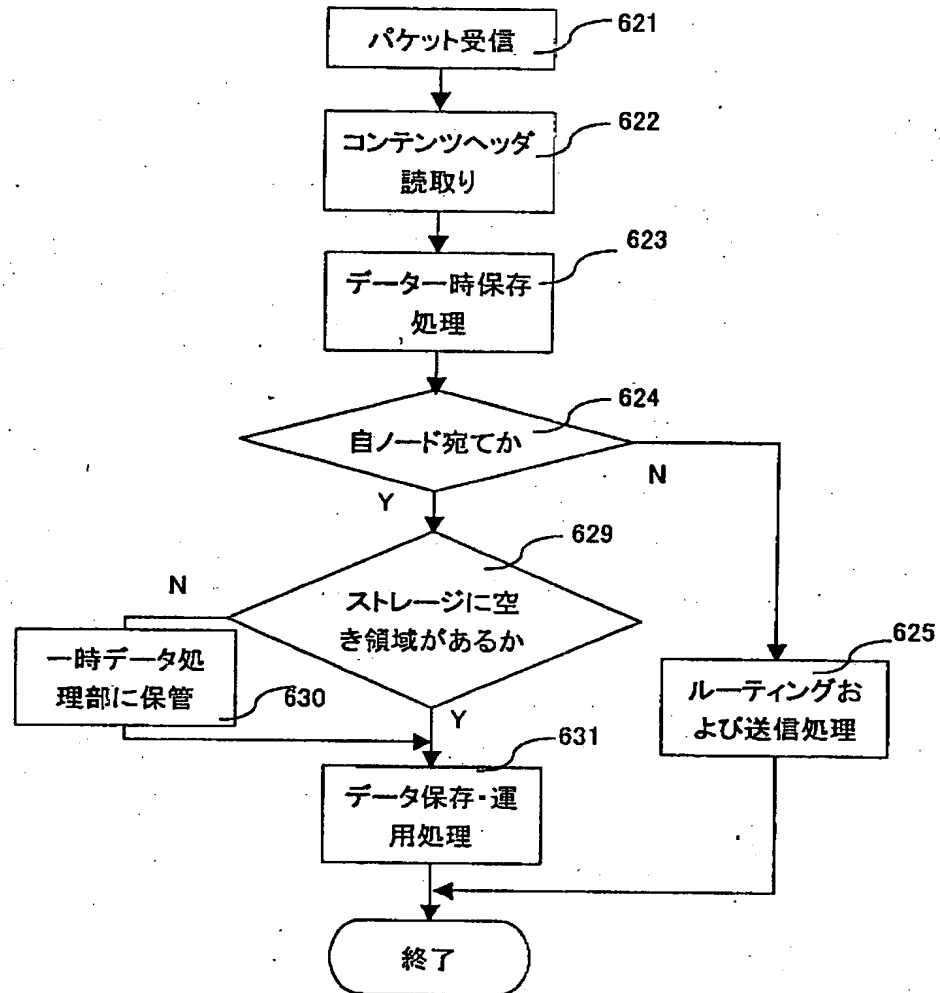
【図 15】

図 15



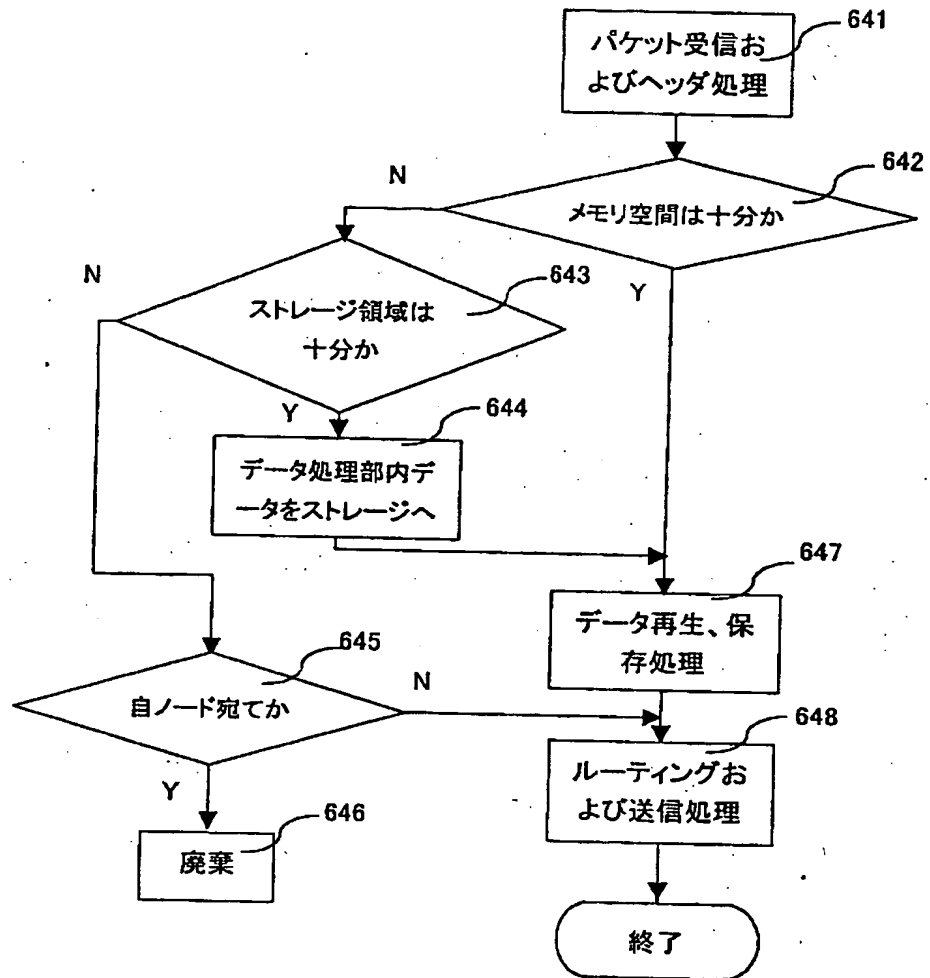
【図16】

図 16



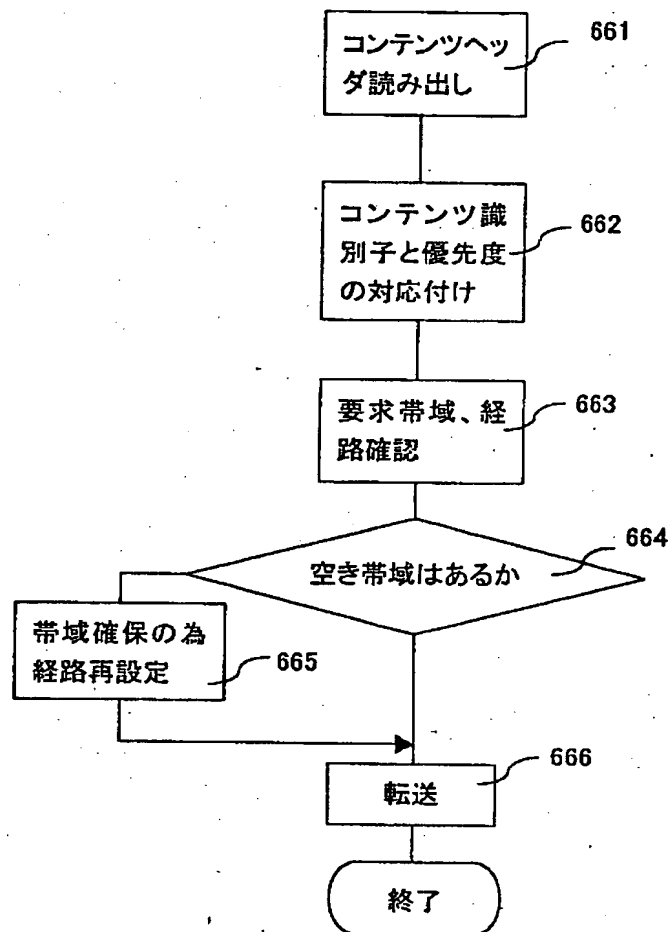
【図 17】

図 17



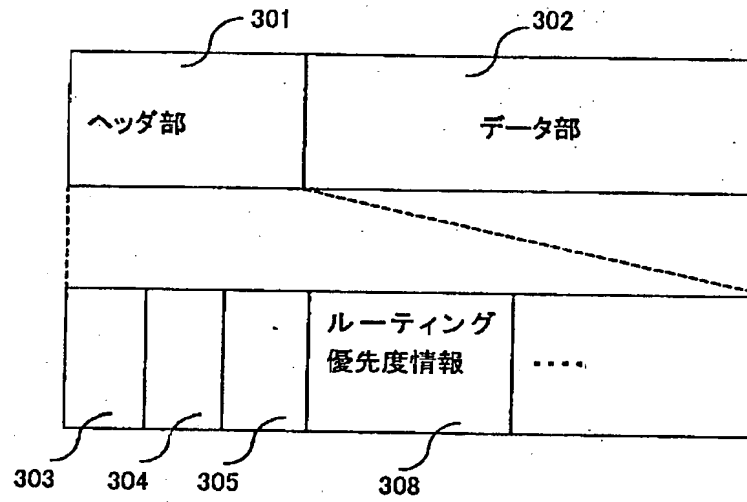
【図 1. 8】

図 18



【図 19】

図 19



【図 20】

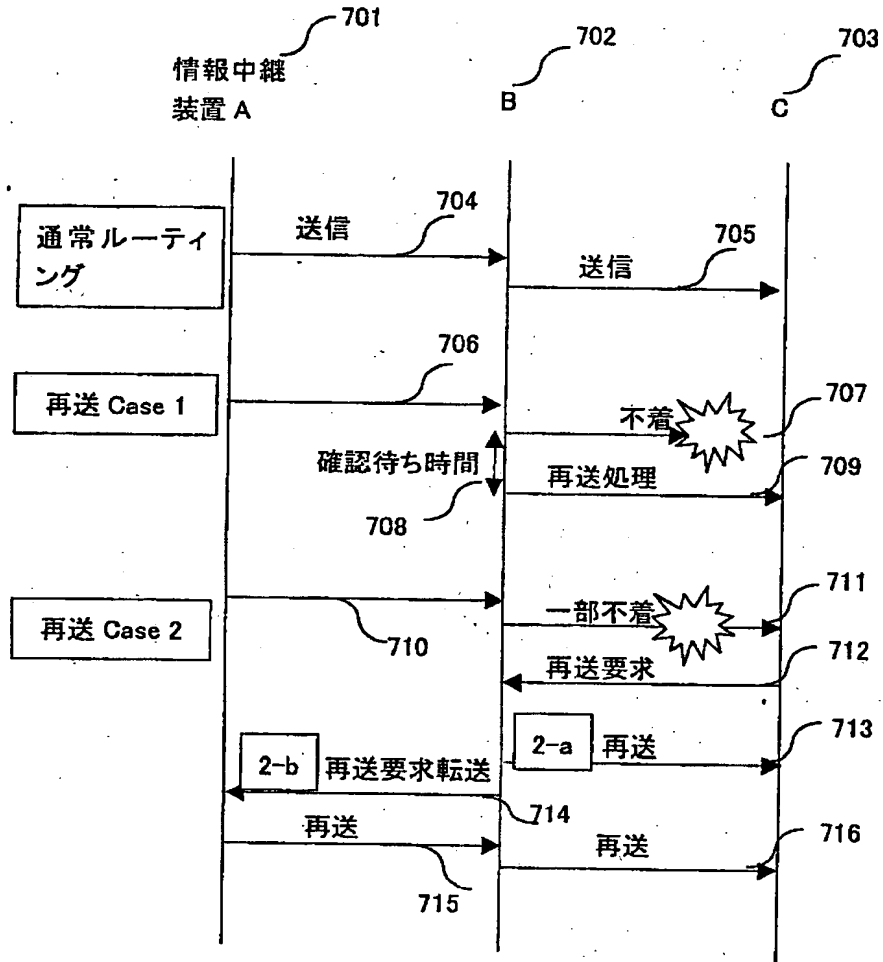
図 20

宛先アドレス	Next hop アドレス	利用可能帯域
A	a	20Gb
B	b	50Gb
⋮	⋮	⋮

Diagram illustrating a table structure. The table has three columns: 宛先アドレス (Destination Address), Next hop アドレス (Next hop Address), and 利用可能帯域 (Available Bandwidth). The rows are labeled A, B, and ⋮ (vertical ellipsis). The table is divided into three main sections: 401 (Destination Address), 410 (Next hop Address), and 402 (Available Bandwidth). The rows are further divided into sub-sections: 411 (A), 414 (B), and 420 (⋮). The sub-sections are further divided into: 404 (A), 421 (B), and 424 (⋮).

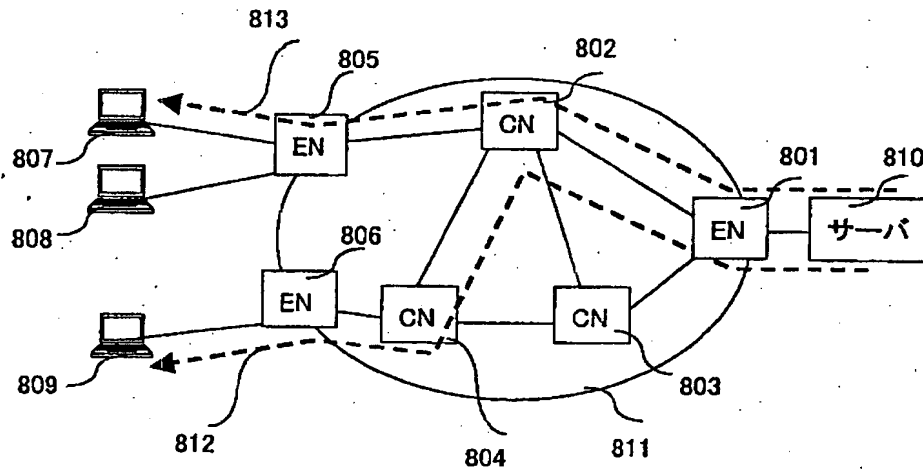
【図 21】

図 21



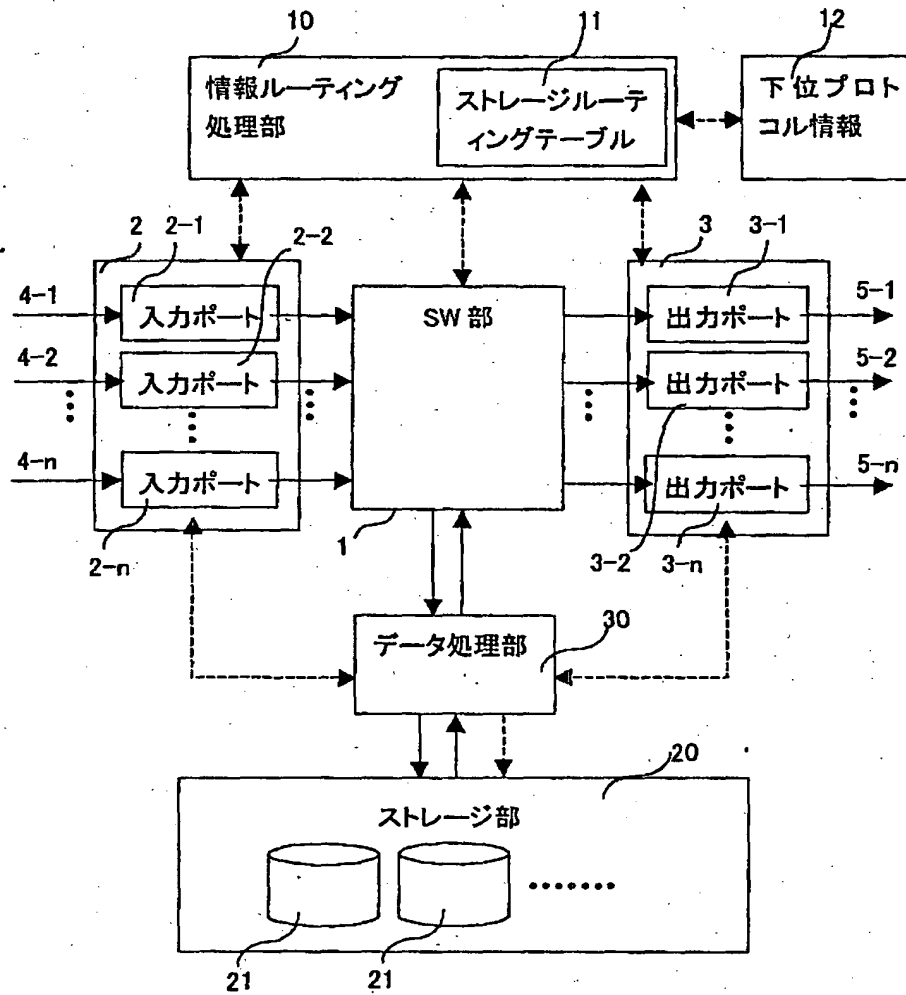
【図 2 2】

図 22



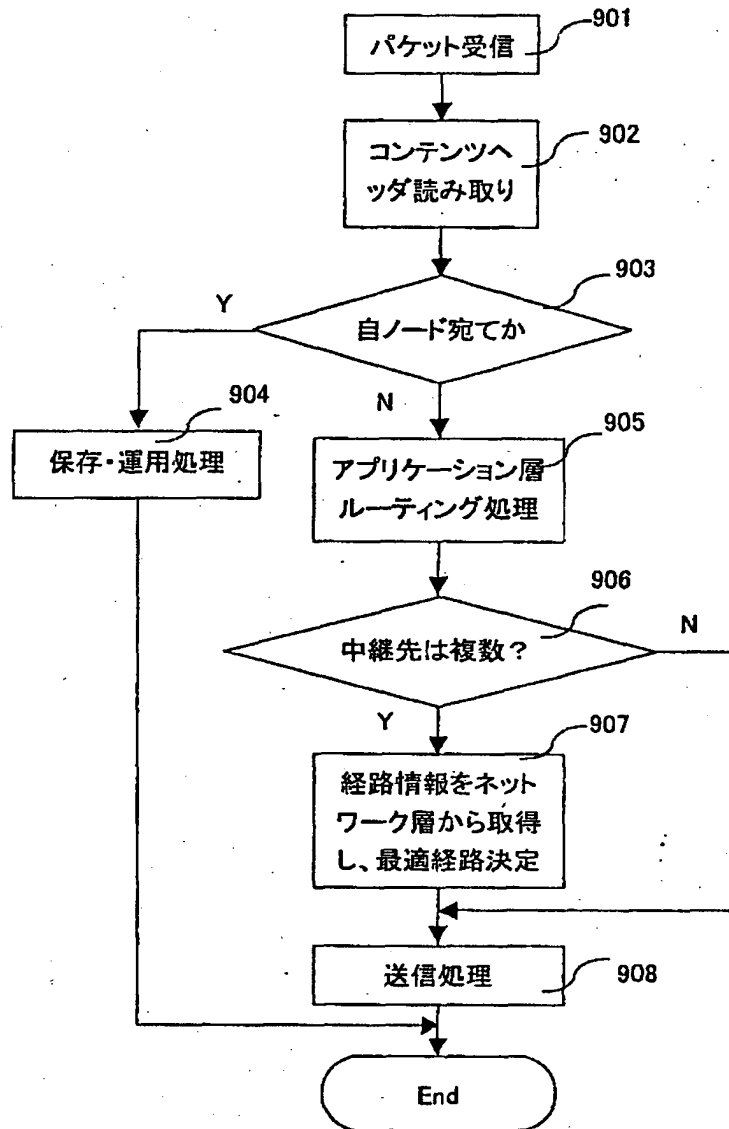
【図 23】

図 23



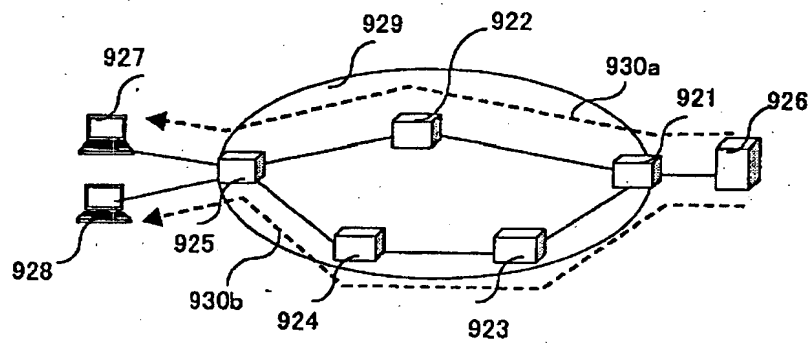
【図 24】

図 24



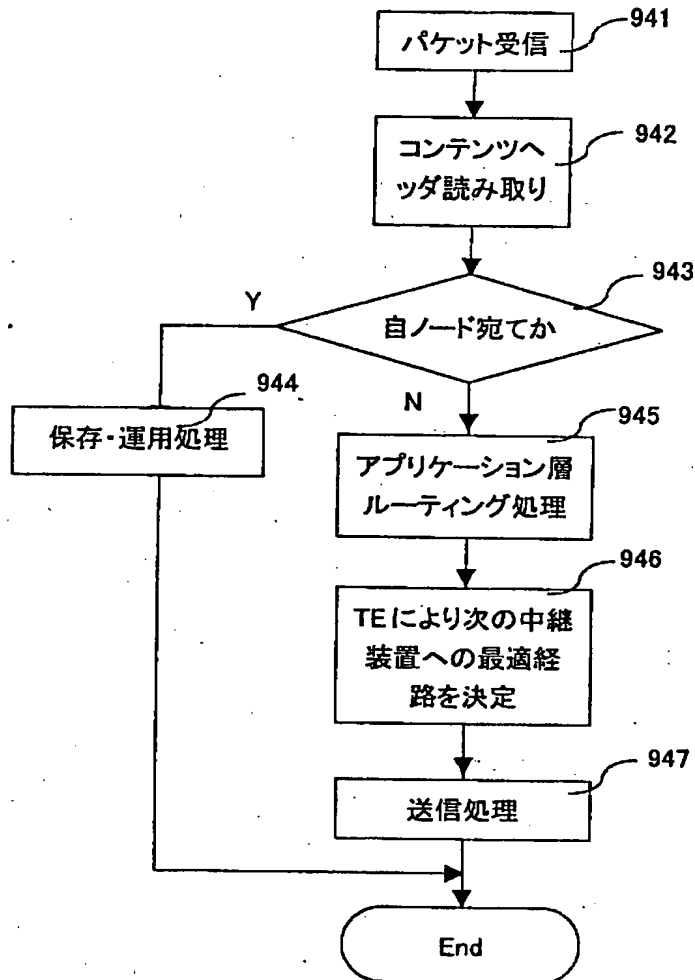
【図 2 5】

図 25



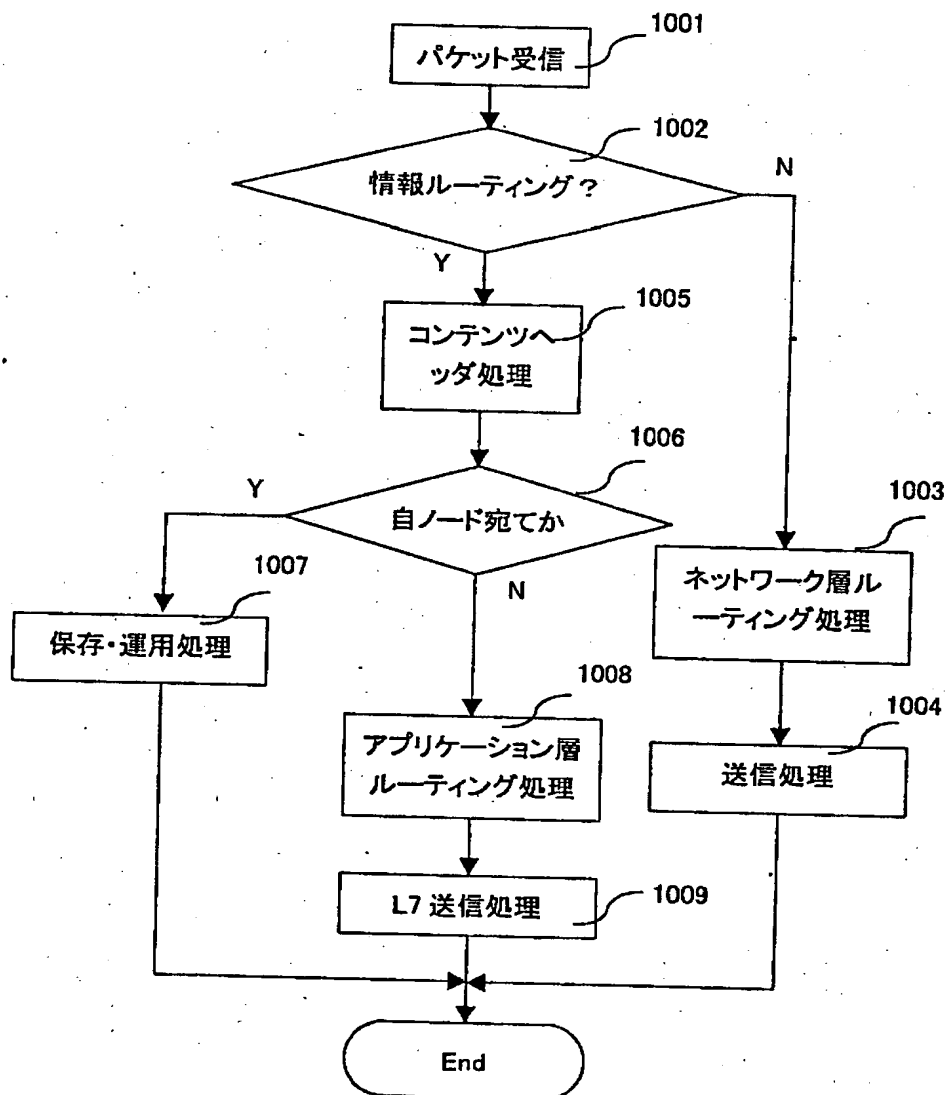
【図 2 6】

図 26



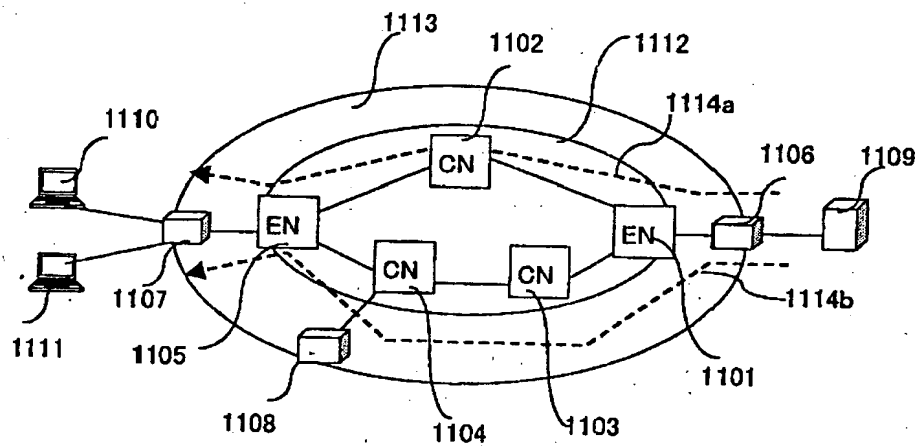
【図 27】

図 27



【图 28】

图 28



【書類名】 要約書

【要約】

【課題】大容量データの配信時における既存ネットワークでの帯域不足による輻輳を回避する手段を提供する。また、大容量データをネットワーク上で確実に送信する為の手段を提供する。

【解決手段】ルーティング処理部10は、ストレージアドレスを含む経路情報に基づいて受信データパケットの送出先を、送信部3またはデータ処理部30から選択し、スイッチ部1に対して切替を命令する手段と、ストレージ部20は、受信データを記憶する手段とを有し、スイッチ部1は、前記制御信号に基づき経路の切替を行う手段を有し、データ処理部30は、データ属性に基づいて、データを格納もしくは送信する手段を有し、送信部3は、データ処理部30からの制御信号に従いパケットのヘッダ処理を行う手段と、データパケットを隣接中継装置へ転送する手段と、を備えることを特徴とする。ネットワークストレージアドレスを用いて表されるストレージルーティングテーブルによりルーティングを行うことを特徴とし、また、データ処理部30およびストレージ部20はヘッダ情報をもとにデータパケットを識別し、少なくとも次段のNSAへの転送が完了するまで受信したデータパケットの情報を装置内に保存しておく手段を有することを特徴とすることで、確実な転送手段を得ると共にネットワーク帯域の有効活用を図る。

【効果】従来のネットワーク技術では難しかった帯域の有効活用が可能となる。

【選択図】 図1

認定・付加情報

特許出願の番号	特願2002-302443
受付番号	50201559799
書類名	特許願
担当官	第八担当上席 0097
作成日	平成14年10月18日

<認定情報・付加情報>
【提出日】

平成14年10月17日

出 願 人 履 歴 情 報

識別番号

[000005108]

1. 変更年月日	1990年 8月31日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台4丁目6番地
氏 名	株式会社日立製作所